How important are human capital, physical capital and total factor productivity for determining state economic growth in the United States: 1840-2000?

Turner, Chad and Tamura, Robert and Mulholland, Sean

Texas AM University Corpus Christi, Clemson University, Stonehill College

11 February 2008

How Important are Human Capital, Physical Capital and Total Factor Productivity for Determining State Economic Growth in the United States, 1840-2000?*

Chad Turner[†], Robert Tamura[‡], Sean E. Mulholland[§]

July 2011

Abstract

This paper introduces new data on state-level physical capital by sector and land in the farm sector for the states of the United States from 1840 to 2000. These data are incorporated into aggregate accounting exercises, with the aim of comparing cross-state results to those found in cross-country samples. Our aggregate results agree closely with the cross-country literature: input accumulation accounts for most of output growth, but variation in the level and growth of TFP accounts for most of the variation in the level and growth of output per worker.

*We thank the workshop participants at Pepperdine University, Midwest Macroeconomics Meetings at the University of Iowa, Nicholls State University, and Xavier University. We have benefited from discussions with Curtis Simon, Gerald P. Dwyer, Mark Fisher, Paula Tkac, and Kevin M. Murphy. We particularly thank Scott Baier and Todd Schoellman for all their generous suggestions. All ideas and remaining errors are the authors' and do not necessarily represent those of the Atlanta Federal Reserve Bank or the Federal Reserve System.

[†]Department of Finance, Economics & Decision Sciences, Texas A&M University: Corpus Christi; chad.turner@tamucc.edu

[‡]Department of Economics, Clemson University and Federal Reserve Bank of Atlanta; rtamura@clemson.edu

[§]Department of Economics, Stonehill College; smulholland@stonehill.edu.

1 Introduction

Empirical research seeking to explain cross-country output per worker differences has expanded greatly since the introduction of new data, particularly the Penn World Tables (Heston, Summers and Aten, 2009) and Barro and Lee's schooling data (Barro and Lee, 2001). Recent advances have been based mostly on development and growth accounting. With some presumed knowledge of a common production function, it is possible to decompose countries' output per worker levels and growth experiences into the relative contributions of measured inputs and the efficiency with which those inputs are used, commonly called total factor productivity (TFP). Accounting exercises seek to understand what proportion of output variation, output growth, and output growth variation can be understood through input variation, input growth, and input growth variation, and what proportion is left to TFP. The typical answer has been that most output growth can be accounted for by input growth, but that the variation in output levels and growth rates is mostly accounted for by variation in TFP levels and growth rates.[1] These results have spurred a search for a theory of TFP.[2]

Our contribution is to apply the methodology of cross-country development and growth accounting to a panel data set covering the states of the United States from 1840 to 2000. Using states as a laboratory offers three main advantages when compared to countries. First, output per worker has converged among states in a way that it has not among countries, allowing us to ask whether the main accounting results hold throughout the convergence process. Second, there are more detailed data available for states than for countries. In particular, sectoral level data are available, making it possible to do sectoral accounting exercises. Finally, to some extent TFP is famously merely a "measure of our ignorance." For the many readers more familiar with the institutions and laws of states than of countries, it will be easier to identify the other factors that are, of necessity, left out of accounting exercises.[3] Throughout, our goal is to use our results to help understand cross-country results, and particularly to see which results extend to states and which do

[1] For example, most of the East Asian miracles' growth experiences can be accounted for through input accumulation, see Young (1995, 2003). However, Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), and Caselli (2005) find that TFP variation accounts for most of output per worker and output per worker growth variation.

[2] This search can be traced back at least to Prescott (1998).

[3] Prior research has identified a number of potentially important factors, including differences in legal systems (Berkowitz and Clay, 2005, 2006), disease burden (Bleakley, 2007, 2010), banking systems (Bodenhorn, 2002), and the legacy of slavery in the South (Mitchener and McLean, 2003). Doubtless numerous other factors apply. The general point is that these factors are more familiar to most readers than the institutional differences between, for example, Mongolia and Bolivia, and in many cases more homogeneous across states.

not.

We construct the panel data set necessary for growth accounting exercises. In particular, we construct original per worker estimates of human capital, physical capital and land use in agriculture at the state level covering 1840-2000. These data cover 28 states for the entire 1840-2000 period, with other states becoming available later (generally, around the time of statehood). The physical capital data are disaggregated into three broad sectors: farm, manufacturing, and non-manufacturing & non-farm (hereafter NMNF). These data for the farm and manufacturing sectors are drawn directly from historical measures of, or proxies for, each state's capital stock, while our measures for the remaining sector requires some imputation. Our measure of land covers the farm sector and includes adjustments to account for the innate productivity differences of land of different types, and different areas of the country. A companion paper describes the construction and robustness of these figures in more detail.[4] When combined with measures of human capital, the labor force, and output by state and sector from Turner, Tamura, Mulholland and Baier (2007), we have sufficient data for our accounting exercises. One major difference in this paper, compared with Turner, Tamura, Mulholland and Baier (2007), hereafter TTMB, arises from our measure of output per worker. The current paper has a consistent measure of real output per worker over the entire 1840 - 2000 period, while TTMB uses a combination of real income per worker (1929-2000) and real output per worker (1840-1920). In addition to desiring consistency, we prefer the output per worker series as the cross-country accounting literature typically utilizes output per worker measures, and thus this choice facilitates comparisons to the cross-country literature. In addition, we better fit observed national capital-output ratios using output per worker.

We first replicate the standard cross-country growth accounting exercises on the aggregate state data. Output per worker converged substantially over the 160 years of our survey. The ratio between the most and least productive states was generally over 5 in the 19th century, and was over 15 in 1850 and again in 1870. It fell to between 3 and 5 in the first half of the 20th century, and was generally less than 3 in the second half of the twentieth century. Hence, we are interested in whether the conclusions concerning the relative importance of TFP hold throughout this period of convergence. We find that they do: variation in the growth of TFP accounts for most of the variation in the growth rates of states.

Our paper is relevant to both the literature that aims to understand the growth ex-

[4]See Turner, Tamura, Schoellman and Mulholland (2011). This paper details the construction of the capital and land data for the period 1850-2000. We detail below how the 1840 values of capital and land are produced from these data.

perience of the United States specifically, and that which aims to understand the growth experience across countries. Jorgenson (2005) is a recent example that includes an overview of the relevant literature for the United States. We are aware that our methodology is simpler than many of the papers in this portion of the literature, and concede these papers include more detailed accounting and often include more factors than our study. That being said, the reason we adhere to our relatively simple methodology is to make as direct a comparison as possible to the cross-country literature. We also note, that relative to both sets of literature, we offer an accounting exercise that contains a longer dataset than typically available.

The paper proceeds as follows. Section 2 describes the creation and attributes of our panel data set briefly, with further description available in the companion paper Turner, Tamura, Schoellman and Mulholland (2011), hereafter TTSM. Section 3 performs the aggregate accounting exercises. Section 4 contains robustness checks, with alternative time periods, alternative input measures and sectoral accounting exercises. The results from both the growth accounting and the variance decomposition of growth rates are robust to different input measures, different time periods and across sectors of the economy. Section 5 concludes.

2 Measurement of Inputs and Output

The first contribution of this paper is to provide original estimates of physical capital and land for each state. Our estimates of physical capital date back to 1840 for 28 states, with coverage for the last states (Alaska and Hawaii) beginning in 1950. Our estimates of land in the farm sector date back to 1840 for 28 states, with coverage for the last state (Alaska and Hawaii) beginning in 1900. We do not have reliable figures of land in Washington, D.C.[5]

Our estimates of total physical capital are decomposed into estimates for the farm and manufacturing sectors, as well as estimates for the NMNF sector. Our estimates for the farm and manufacturing sectors are based on historical state-level observations on capital or proxies for capital, derived primarily from the Censuses of Agriculture and Manufactures. Our estimates for the NMNF sector involve more imputation. Our estimates of land cover the value of land in the farm sector. Following the convention in much of the literature, we will generally ignore the role of land when doing aggregate or manufacturing sector

[5] *The Census of Agriculture* stopped enumerating it separately in 1935 due to relatively low levels of farm activity.

accounting exercises, but later include land in robustness checks.[6]

Below we introduce our data sources and measurement procedures for capital and land.[7] When robustness checks are suggested here, the details are available in our companion paper, TTSM. While data for human capital, and the labor force were previously collected and presented in TTMB (2007), we overview the main features of those data as well.

2.1 Physical Capital in the Farm Sector

We use three primary data sources for measuring physical capital in the farm sector. For the 1850-1959 period we rely on the *Census of Agriculture*, with the data often tabulated in a more convenient form in the *Statistical Abstract of the United States*. For 1960-2000 we use the annual data provided by the United States Department of Agricutlure (USDA) in their Farm Balance Sheets. Finally, some special data we need are available only in special reports to the Census of Agriculture, titled variously the *Farm Finance Survey* (following on the 1969 and 1978 Censuses of Agriculture) or the *Agricultural Economics and Land Ownership Survey* (1987 and 1997 Censuses).[8]

With these data sources it is possible to construct a long time series of state-level observations on land & building value, implements/machinery/equipment value, and livestock value. Additionally, the Farm Balance Sheets include data on crop inventory value for 1960 onward. Our measure of farm sector capital is closely related to the sum of these series, but we face three difficulties. First, by adding these series we include crop inventories for some years but not others. Because crop inventories are generally a small share of the capital stock, we choose to include all available data without adjusting for inclusion of crop inventories.

Second, land & building value actually measures the value of land, farm buildings, and farm dwellings. We account for land separately and do not want to include farm dwellings in the farm capital stock. We remove the value of land and farm dwellings in two steps. The 1900-1940 Censuses and the special reports to the 1969, 1978, 1987, and 1997 Censuses

[6] In one robustness check we conduct growth accounting for farming with land as an input, and in another robustness check we construct an alternative aggregate input measure which includes land for the aggregate growth accounting exercise. In both cases we also repeat the variance decomposition exercises on growth rates.

[7] See TTSM for further detail.

[8] For the 28 states that we observe in 1840, the 1840 farm capital and land valuations come from back projections for each state. We have farm capital and land valuations from 1850-2000, so we regressed log(land) against time and time squared for each state, and used the predicted value from each state regression for the initial 1840 value. For farm capital we regressed log(farm capital) on log(land), time and time squared for each state. We then used the predicted value in 1840 for the 28 states we observe in 1840.

enumerate land value separate from building/dwelling value, so in these Censuses land is already removed. For other years, we estimate building/dwelling value as:

$$buildings/dwellings = land/buildings/dwellings * \frac{\widehat{buildings/dwellings}}{land/buildings/dwellings}$$

where we approximate buildings/dwellings share of the total by linearly interpolating from available data between 1940 and 1969, and projecting the 1900 values back to 1850. These shares do fluctuate systematically over time as the relative value of land changes; for example, land was relatively expensive in the 1970s, leading to a lower buildings/dwellings share in 1978.

Separate data on the value of farm dwellings and buildings are even scarcer: an explicit breakdown of farm dwellings versus other farm buildings was asked in the 1930 Census and special reports to the 1987 and 1997 Censuses. The share of farm dwellings in total building value appears relatively constant in these years, so we estimate farm building value in other years as:

$$buildings = buildings/dwellings * \frac{\widehat{buildings}}{buildings/dwellings}$$

We combine our estimated farm buildings value with implements/machinery/equipment, livestock, and crop inventory (when available) to yield our total agricultural capital stock.

Some imputations are necessary for Alaska, Hawaii, and Washington, D.C., which have less data available. We aggregate our results and compare them with BEA estimates of the national physical capital stock in the farm sector for 1947 onward; we overestimate their capital figures in early years, agree fairly closely in the 1960s and 1970s, and underestimate starting around 1980.[9]

2.2 Physical Capital in the Manufacturing Sector

We use two primary data sources for measuring physical capital in the manufacturing sector. We use *Census of Manufactures* data for the entire 1850-1997 data, but complement it with more frequent data from the Annual Survey of Manufactures when possible. Data from both sources are often tabulated in the *Statistical Abstract of the United States*.

The Census of Manufactures contains the capital stock by state for the period 1850-

[9]For the subsequent empirical analyses, we exclude Washington, D.C.

1919.[10] Their measure of the capital stock includes both fixed capital (capital in the modern sense) as well as working capital (cash on hands, bills receivable, and so on). In 1890 and 1900 they asked respondents to distinguish between the two, giving us a direct measure of (fixed) capital. The relative proportion of fixed to working capital seems relatively constant across states and years, so we use the relationship to estimate capital for the entire 1850-1919 period as:

$$fixedcapital = capital * \frac{\widehat{fixedcapital}}{capital}$$

where the proportion of capital that is fixed is estimated using the 1890 and 1900 data.

The 1919 Census was the last to ask respondents about the value of their capital until the 1977 Census of Manufactures; from 1977-1997, we have information on the gross book value of capital by state. The primary difficulty in the manufacturing sector is constructing reliable measures of capital in the intervening years. We have two sources of data. First, information on new capital expenditures (investment) was asked in 1939, 1947, and annually for most states beginning in 1952. This data obviously leaves substantial gaps. In the interceding years the Census Bureau collected various measures of horsepower in use by state. We construct alternative measures of the capital stock using this data, but ultimately decide against using it here because of outside evidence that the relationship between the dollar value of capital and horsepower in use is not stable over this period (Tostlebe 1954).

Instead, our baseline approach is to construct an estimate of the 1939 capital stock using $K_{1939} = I_{1939}/\delta$.[11] We then interpolate (the log of) investment for missing years between 1939 and 1952, and use a perpetual inventory method to construct capital stock for 1940-1972,

$$K_{t+1} = (1 - \delta)K_t + I_t,$$

with $\delta = 0.06$ as in Caselli (2005). Finally, we interpolate the capital stock between 1919 and 1939, and then between later Census years as necessary.

We again make some corrections for the lesser data availability in Alaska, Hawaii, and

[10] For the 28 states that we observe in 1840 we again used back projection to get our initial manufacturing capital stock. Using data from 1850 onward, for each state, we regressed log(manufacturing capital) against time and time squared, log(land value) and log(farm capital). We then used the predicted value from each regression for the initial manufacturing capital in 1840.

[11] Caselli (2005) uses the Solow steady-state capital stock of $I_{1939}/(\delta + g)$ where g is the average growth rate of investment over subsequent years. We do not use this approach since we lack annual data until the 1950s.

Washington, D.C. We aggregate out state results and check against the BEA estimates of national physical capital in manufacturing from 1947 onward and find that they agree closely throughout the time period.

2.3 Physical Capital in the NMNF Sector

Censuses for the other industries began later and lack measurement of capital stocks suitable for our purposes.[12] Given the lesser capital data availability, we are forced to impute data for the capital stock in these industries. Our basic approach is to use the approximation:

$$K_{it} = \sum_j \frac{K_{ijt}}{Y_{ijt}} Y_{ijt} \approx \sum_j \frac{K_{jt}}{Y_{jt}} Y_{ijt}. \tag{1}$$

where K is capital, Y is output, and i, j, and t subscripts denote states, industries, and years. Our NMNF industries closely mirror those of the BEA data: mining, construction, transportation and public utilities, FIRE, services, and government. We aggregate wholesale and retail trade since data are not available separately prior to 1958. Hence, we use aggregate data on capital-output by industry and the state composition of output by industry to impute each state's capital stock for the NMNF sector.[13] Doing so requires ignoring any state-level variation in capital-intensity within an SIC industry, an assumption previously made by Garofalo and Yamarik (2002) in their work on regional convergence from 1977 to 1996; of course, we have reason to suspect that capital-output ratios may be more similar across states at the end of the 20th century than they were in the 19th. Given the lack of outside data it is difficult for us to circumvent or test this assumption.

We construct the aggregate NMNF industry capital-output ratio using NIPA data, available from 1947-1997. Outside this range we are forced to impute the data, but most industries show no trend. Indeed, Gallman (1986) reports values for several industries from 1840-1900 that closely agree with those observed from 1947-1997 (his Table 4.8). For these industries we use the average $\frac{K_{it}}{Y_{it}}$ observed. For wholesale trade and the government there is a significant trend; we assign the 1947 value to all years before 1947, and the 1997 value to 1998-2000. Our results are robust to allowing for time trends.

From 1963 through 1997, the BEA provides measures of gross state product in each industry for each state. While the BEA does not provide estimates of GSP from 1929-1962

[12] For example, the Census of Mining first collected a proxy for capital, aggregate horsepower, in 1902, while the Census of Finance, Insurance, and Real Estate was not separated from the Census of Construction until 1992.

[13] Since we are focusing on the 1840 - 2000 period, we did not use the seminal work of Jorgenson, Gollop and Fraumeni (1987), which produces sectoral capital measures from 1948 to 1979, inclusive.

or from 1998-2000, they do provide measures of wages and salary disbursements in each industry at the state level (series SA07). We find that gross state product and wages are highly correlated. We fit a relationship between GSP and wages, and use this to predict GSP by industry for 1929-1946 and 1998-2000.[14] The fit is generally quite good; mining has the worst fit,with an R^2 of 0.95. We use actual or imputed GSP by industry and aggregate capital-output ratios by industry to construct the state NMNF capital stock using equation (1) for 1929-2000.

Before 1929, our only data source is GSP for farming, manufacturing, and NMNF.[15] While we experimented with several other methodologies, we ultimately chose to use the simple method described here, as it delivered similar results and required fewer assumptions.[16] We take the average capital-output ratio in the NMNF sector for each state over the 1929 - 2000 period, and apply it to the output measures in the NMNF sector for all years prior to 1929.

2.4 Physical Capital per Worker

We sum our measure of physical capital in the three sectors to arrive at a state-level total physical capital stock. Figure 1a displays the labor force-weighted average physical capital per worker for the United States, as well as the coefficient of variation and figures for the state with the least and most capital per worker. The general process of convergence is

[14]The baseline regression is log-log and controls for state fixed effects.

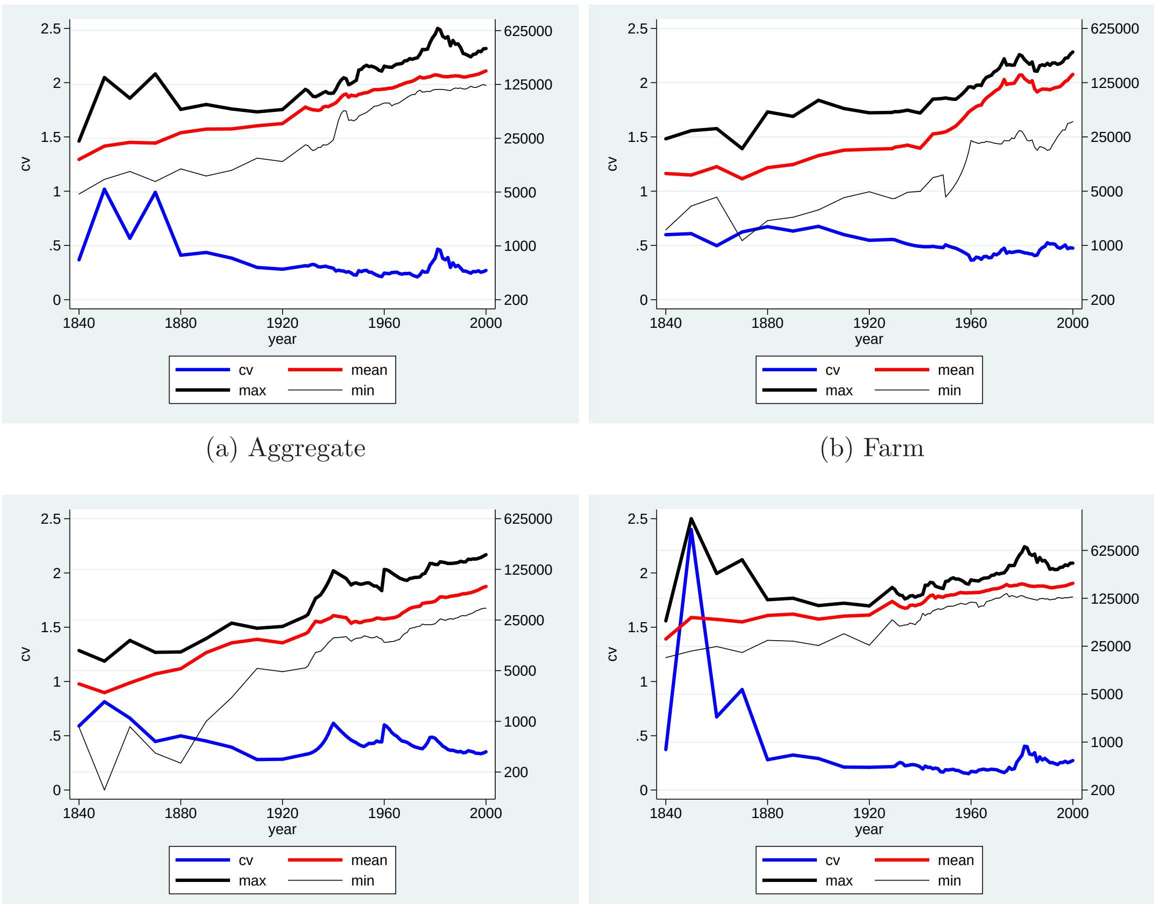
[15]We removed fishing and forestry from agriculture to get farming separately. We did this by constructing, for each state separately, the average share of agriculture that is farming using data from 1929-2000. We then applied these farming shares to agriculture in order to produce a consistent farming series from 1840-2000.

[16]In an earlier version we projected the value-added share of each of the seven industries in the NMNF sector back from 1929 to 1840. Our projection was $\log(gsp_{ijt}) = \beta_0 + \beta_1 * \log(gsp_{it}) + \beta_2 * \log(pop_{ijt}) + \sum_i \alpha_i D_i + \sum_i \gamma_i D_i t$ where pop represents population, the D_i are state dummy variables, and $D_i t$ are state-time interactions. We then constructed capital stocks as:

$$K_{it} = \sum_j \frac{K_{ijt}}{Y_{ijt}} \frac{Y_{ijt}}{Y_{it}} Y_{it} \approx \sum_j \widehat{\frac{K_{jt}}{Y_{jt}}} \widehat{\frac{Y_{ijt}}{Y_{it}}} Y_{it}. \tag{2}$$

where $\widehat{\frac{Y_{ijt}}{Y_{it}}}$ denotes a projection of industry i's share of state j's NMNF production at time t. As a check on our data, we also project the agriculture and manufacturing output share of each state back in time. As discussed in TTSM, we have outside evidence on output for these industries at the state level for 1840 to 1920. Hence, we can use this data as a check on the validity of our projections. We find that our projections deliver a good fit in manufacturing ($R^2 = 0.64$), but not as good a fit in agriculture ($R^2 = 0.25$), driven by a poor fit in Western states. We then constructed capital for 1840-1928 using equation (2). We tried several alternative ways of projecting back NMNF and found that most reasonable alternative specifications yield capital stock measures with similar properties, as measured by the correlation across alternative capital stocks and the average growth rate.

immediately apparent in the aggregate and in all three sectors.



(a) Aggregate

(b) Farm

(c) Manufacturing

(d) NMNF

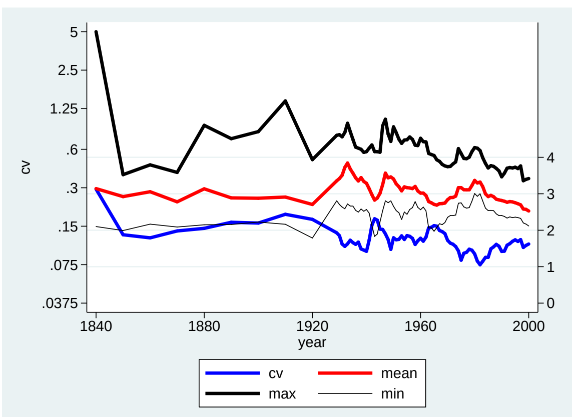
Figure 1: Physical Capital per Worker

Figures 1b and 1c give the capital per worker in the farm and manufacturing sectors. Capital per worker differences are larger and show less convergence in the farm sector than in the manufacturing sector. Not surprisingly, there is not as much dispersion in capital per worker in the NMNF sector. In 1840 the coefficient of variation of physical capital is .43, and rises to a peak of 1.1 in 1870 before dropping to .44 in 1900.[17] From 1910 to the present it has varied from .33 to .44, except for the 1978-1985 period when it varies between .46 and .64.

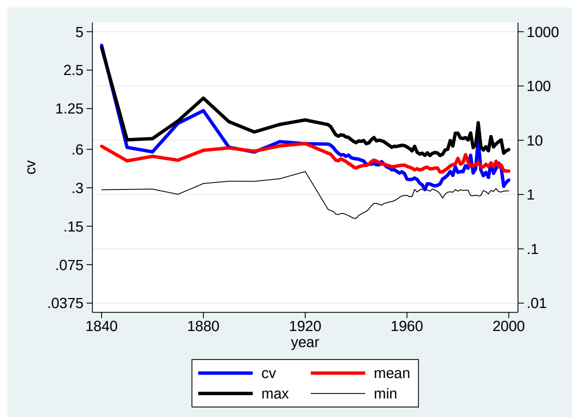
Variations in capital per worker arise from two sources, variations in sectoral shares, and variations in capital-output ratios. Figure 2 presents the data on the capital-output ratio

[17]In 1870 without California, Colorado and Nevada, the coefficient of variation of capital is .48 instead of 1.1.

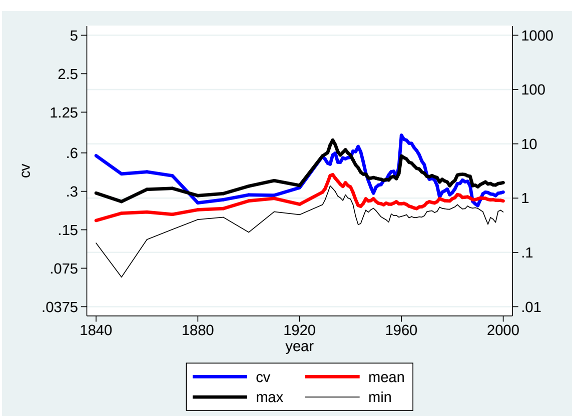
for the aggregate, farm sector, manufacturing sector and the NMNF sector. We present both the average capital-output ratio, weighted by the labor force of the state, as well as the minimum, maximum and coefficient of variation of the capital-output ratio. All graphs in Figure 2 have the same left vertical scale (for the coefficient of variation) and the same right vertical scale (for the capital-output ratios), except for the aggregate physical capital per worker graph. The average aggregate capital output ratio starts at about 3 in 1840, rises to 3.4 in 1929, to 3.6 in 1947, declining to 3 in 1984 and continuing to trend lower to 2.75 by 1995 and 2.5 in 2000. The coefficient of variation of the capital-output ratio begins at .46 in 1840, falls to .24 in 1850, and remains in a narrow range between .17 and .27. [18]



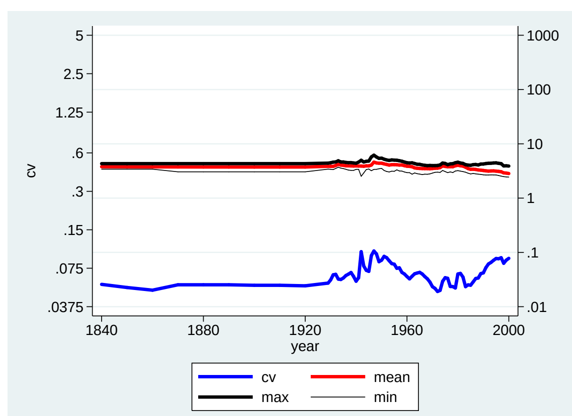
(a) Aggregate



(b) Farm



(c) Manufacturing



(d) NMNF

Figure 2: Physical Capital - Output ratio

Table 1 below illustrates the change in dispersion of physical capital per worker by

[18] In 1840, the coefficient of variation of the capital-output ratio without Iowa, is .28 instead of .46.

sector in farming and manufacturing with our improved data.[19] Under the assumption of a common capital-output ratio in each sector for any state, the variance of the log of physical capital per worker by sector would be equal to the variance of the log of output per worker by sector by definition. We report the variance of the log of physical capital per worker in farming and manufacturing, by decade, from 1840-2000. Except for the 1840, 1860 and 1920 observations in farming, and the 1860, 1870, and 1890-1930 observations in manufacturing, the standard deviation of log physical capital per worker exceeds the standard deviation of log output per worker in farming and manufacturing. Thus our new measures of real physical capital per worker display greater cross state variation than our measures of cross state variation in real output per worker.

Table 1: Variation in Sectoral log output per worker and log physical capital per worker

		Farming		Manufacturing	
	Obs	σ_{lnk}	σ_{lny}	σ_{lnk}	σ_{lny}
1840	28	0.6990	0.7841	0.6520	0.4505
1850	34	0.4977	0.4866	0.9485	0.8642
1860	38	0.4392	0.4418	0.5650	0.5716
1870	45	0.7614	0.7102	0.5531	0.6757
1880	45	0.7053	0.5131	0.7511	0.7373
1890	45	0.7015	0.4795	0.7057	0.7059
1900	47	0.7377	0.6118	0.4996	0.5058
1910	47	0.7038	0.6961	0.3160	0.3246
1920	48	0.5855	0.7274	0.3141	0.3570
1930	48	0.6779	0.6270	0.3762	0.4084
1940	48	0.6264	0.6119	0.4993	0.2690
1950	50	0.6075	0.5250	0.3449	0.2051
1960	50	0.3799	0.3364	0.4386	0.2294
1970	50	0.5159	0.2910	0.3648	0.2046
1980	50	0.5554	0.4147	0.3718	0.2012
1990	50	0.6536	0.4507	0.2972	0.2101
2000	50	0.5567	0.5243	0.2995	0.1745

[19] In the previous version of this paper, we used the national capital-output ratio for agriculture, manufacturing and services respectively to produce state estimates of physical capital by sector and aggregate. With the use of the censuses of agriculture and manufacturing, we were able to improve on this procedure. In particular we now have state specific capital-output ratios in farming and manufacturing.

2.5 Land per Worker

A second contribution of this paper is to develop a quality-adjusted series for land in the farm sector by state for the 1840-2000 period.[20] Our land quality adjustment can be understood as a conceptual modification of the labor quality (human capital) adjustment introduced by Bils and Klenow (2000). The Bils and Klenow methodology measures the observable characteristics of workers (such as their years of schooling) and then values those characteristics at their market rate (the Mincerian return to schooling). Similarly, we measure the observable characteristics of land (by dividing it into irrigated improved land, non-irrigated improved land, and unimproved land) and value different types of land at their market rate (given by the spot rental price for that type of land in that state in 2009).

Our approach offers two improvements over a series that measures only the quantity or acreage of land. First, in the cross-section, we are able to account for differences in land quality. The United States is geographically diverse, and the land rental prices we use reflect this diversity. The price to rent an acre of land ranges from $2.40 for New Mexico pastureland to $350 for irrigated California cropland. Even within a fixed category, the price to rent an acre of non-irrigated cropland varies from $12 (Wyoming) to $175 (Iowa). Second, in the time series, we are able to account for investments in land in the form of improving (clearing) and irrigating it, both of which represent major agricultural investments.[21]

Our primary data source for the quantity and type of land is *The Census of Agriculture*, which has asked about the acreage of different types of farmland. These data first allow us to distinguish between improved land that is available for crops, and unimproved land, which is not. We further divide improved land into irrigated and non-irrigated land.

As agriculture spread to the arid Western states, the Census Bureau took a special interest in irrigation, and began enumerating the quantity of irrigated land in Western states with the 1890 Census. They enumerated total acres of irrigated land in Western states for 1890-1920 and 1945 onward. They also enumerated acres of irrigated cropland for 1930-1940 and 1950-1959. Generally, most irrigated land was used for crops, but there are a couple of exceptions (notably Nevada, where half of irrigated land was not used for crops). We use the overlap between irrigated cropland and total irrigated land to measure the (relatively stable) fraction of irrigated land that is cropland. We then measure irrigated

[20] We use the land input in the growth accounting and variance decomposition of growth rates in the farming sector. Also we produce a new measure of aggregate inputs and conduct growth accounting and variance decomposition of growth rates of the aggregate state output per worker.

[21] In an earlier version of this paper, we only used the acreage of arable land.

cropland either directly (when possible) or using total irrigated land and the expected ratio of that land that is irrigated cropland. We do not distinguish between irrigated and non-irrigated unimproved land, since the rental price data we use do not distinguish between them. The fraction of land irrigated is fairly low in most states in their first enumerated year; we assume that no land was irrigated prior to that year.

Irrigation is less important in Eastern states, and enumeration of irrigated land there began only in 1940. Again, the proportion of land that is irrigated is low in 1940, so we assume no land was irrigated prior. Further, the only data collected in Eastern states was irrigated land, never irrigated cropland. We assume that all irrigated land in Eastern states is cropland.

We measure the relative productivity of different acres of land using their cash rental price. We prefer cash rental prices because alternative measures (particularly land values) are affected by non-agricultural uses of land and the expected future value of the land, rather than the current, one-year agricultural value. The USDA produces an *Agricultural Cash Rents*, z, report that gives the rental price for our three types of land, j, in different states, i for that year, t, z_{it}^{j}. They do not report rental prices for all types of land, j, z_{it}^{j}, in all states, i, but the report has become more comprehensive over time; we use the most recent (August 2009) report for this reason. When necessary, we approximate some cash rental prices with the rental price of the same type of land using regional averages or similar states.[22] We combine our measure of quantity of land, by type, per farm worker, λ_{it}^{j} from *The Census of Agriculture* with the estimates of the rental price per acre of land, by type, z_{it}^{j}, from the USDA. Thus our estimate of land input per farm worker is given by:

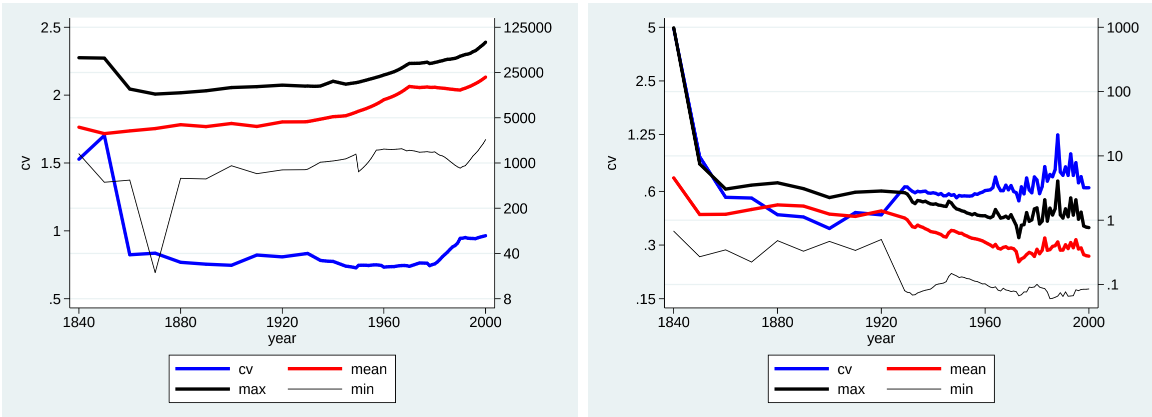
$$\mathcal{L}_{it} = \lambda_{it}^{irr} z_{i,2009}^{irr} + \lambda_{it}^{nonirr} z_{i,2009}^{nonirr} + \lambda_{it}^{unimp} z_{i,2009}^{unimp},$$

where λ_{it}^{j} is the acreage per farm worker of each type j, in each state, i, in year t.

Figure 3a gives descriptive statistics for land per farm worker over time, and figure **??** gives the land to farm output ratio, $\mathcal{L}_{it}/y_{it}^{farm}$. Land per worker has grown throughout the period. Land per worker also shows a fair amount of convergence across states, particularly before 1960, with divergence since 1960. In 1840 the coefficient of variation of land per worker in farming was 1.53, rising to 1.70 in 1850 before dropping by more than half to .82 in 1860.[23] The average value of the coefficient of variation of land per worker from

[22]See TTSM for additional details.

[23]As with physical capital per worker in farming, the 1840 land per worker in farming distribution is strongly affected by Iowa. Ignoring Iowa produces an 1840 coefficient of variation of land per worker in farming of 0.78, about half of the 1.53 value in the figure. The even higher value in 1850 is driven by California. Ignoring California in 1850 produces a coefficient of variation of land per worker in farming of



(a) Land

(b) Land-Farm output ratio

Figure 3: Land per worker in Farm Sector

1860-1959 is 0.770. From 1960 to 2000, there is an increase from .73 in 1960 to .96 in 2000. Thus there is more heterogeneity in land per worker in 2000, than there was in 1840 (excluding Iowa), 1850 (excluding California), and in 1860 for the entire country!

2.6 Human Capital per Worker

Turner, Tamura, Mulholland and Baier (2007) use information on school enrollment rates by state to develop estimates of the average education of the state work force by year, similar to the approach in Barro and Lee (1993), Baier, Dwyer and Tamura (2006) and Tamura, Dwyer, Devereux and Baier (2011). Information on school attainment is transformed into human capital using the log-linear approach pioneered in Bils and Klenow (2000). Log-human capital is assumed to be linear in schooling:

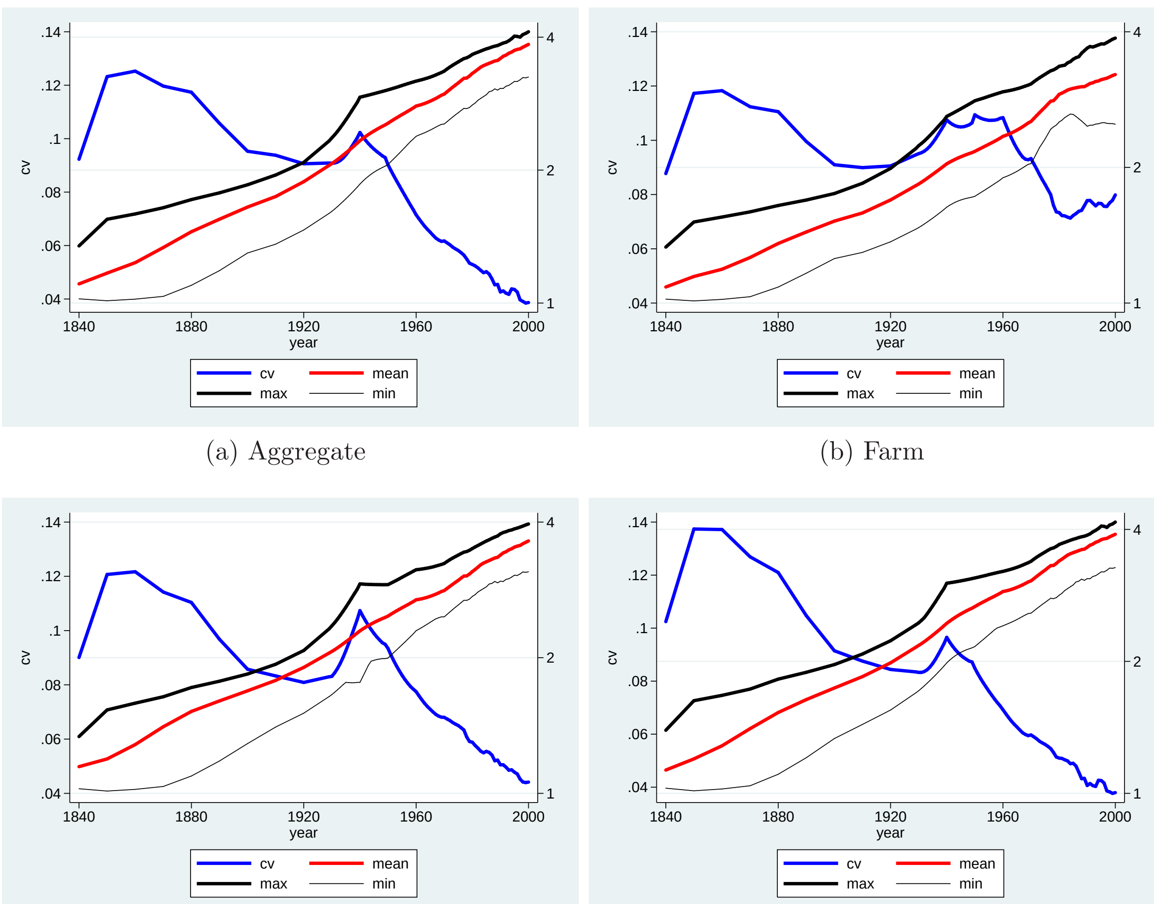
$$\log(h_{it}) = \varphi_P P_{it} + \varphi_S S_{it} + \varphi_T T_{it} \tag{3}$$

where P_{it} represents years of primary schooling, S_{it} years of secondary schooling, T_{it} years of tertiary schooling. In contrast to Hall and Jones (1999) we assume a constant rate of return to schooling, t $\varphi_P = \varphi_S = \varphi_T = 0.10$. These values represent the absence of diminishing returns to schooling in the US found in TTMB (2007) and Card and Krueger (1992).

We also require human capital data by sector. We depart from the assumption of Caselli (2005) that workers in the farm sector have no education. Instead, we use the 1940-2000

0.58, barely one third of the observed value of 1.70!

U.S. Censuses, drawn from the Integrated Public Use Microdata Series system, Ruggles, et. al. (2010). The dataset includes workers' industries recoded into a common system (IND1950); we use this common coding system to improve comparability. For each state we compute the average relative schooling of farming workers to manufacturing workers and the average relative schooling of NMNF workers to manufacturing workers over the 1940-2000 period. For years before 1940 we assign these state specific averages to compute the relative education in farming to manufacturing and NMNF to manufacturing. We then use the accounting identity of average schooling in the state to pin down the schooling level of manufacturing workers. Thus we measure schooling by sector, and construct state sector human capital using equation (3).



(a) Aggregate

(b) Farm

(c) Manufacturing

(d) NMNF

Figure 4: Human Capital Per Worker

Figure 4 displays our human capital results. They show the general increase and convergence in human capital throughout this period, at an aggregate level and for the farm, manufacturing, and NMNF sectors. Human capital in the farm sector begins to diverge around 1980, when an influx of low skilled immigrants into California, Florida, and a few other states lowers the average school attainment in those states.

2.7 Real Output per Worker

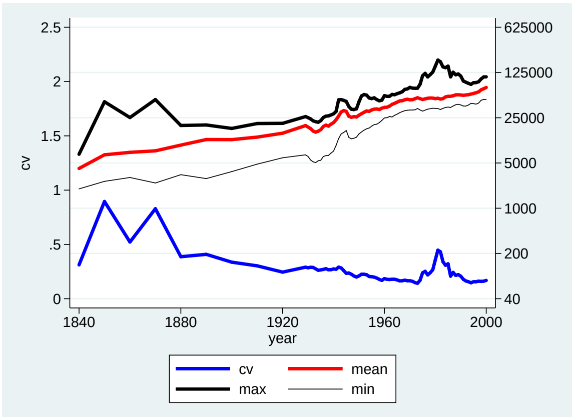
Annual data on output by state are available beginning in 1929 from the Bureau of Economic Analysis. We use Easterlin (1960a, 1960b) estimates of output by state for 1840, 1880, 1900, and 1920, and TTMB measures of state output for 1850, 1860, 1870, 1890, and 1910. The TTMB estimates are based primarily on agricultural and manufacturing output as recorded in the Censuses, separately enumerated mining output for minerals-intensive regions, and imputation of Easterlin's data for other industries. Our output data from 1929-2000 is gross state product, which differs from the real income used in TTMB (2007).[24] Also it is now consistent with cross-country data, which is typically real gross domestic product.

We need information on the size of the labor force and prices to transform nominal output into real output per worker. State labor force measurements in TTMB (2007) draw on *Historical Statistics of the United States*, various census issues, and work done by Weiss (1999) to correct the 19th century census estimates for rural undercounts. Nominal values were converted to real values using information on both annual national price level variation and less frequent observations of interregional price variation.[25]

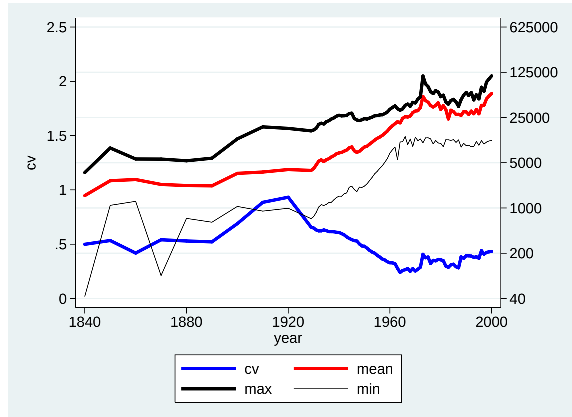
Figure 5a gives the average output per worker for the United States, as well as the coefficient of variation and the output per worker for the richest and poorest state in each year. As with inputs per worker, the general trend has been toward convergence over time. Figures 5b and 5d show that output in the farm and manufacturing sectors both grew over time. However, output per worker in the farm sector is volatile. There is growth per farm worker, and some tendency to converge from 1840-1960. However since 1960 there has been increasing divergence in output per farm worker.

[24]The new series is now consistent, real gross state product, throughout the time period 1840-2000, as opposed to a spliced series contained in TTMB (2007).

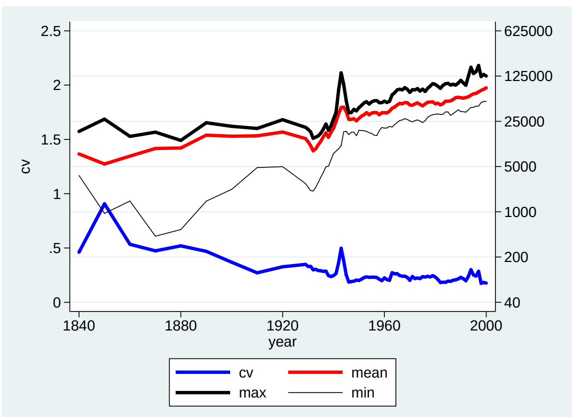
[25]Data on national price level comes from Gordon (1999) for 1875 – 2000, while data prior to 1875 is from *Historical Statistics of the United States*. Regional prices combine information from Mitchener and McLean (1999), Williamson and Linder (1980) and Berry, Fording, and Hanson (2000), which obtains complete coverage except that the Pacific and Mountain regions only have price data beginning in 1880; earlier observations are normalized using national price levels.



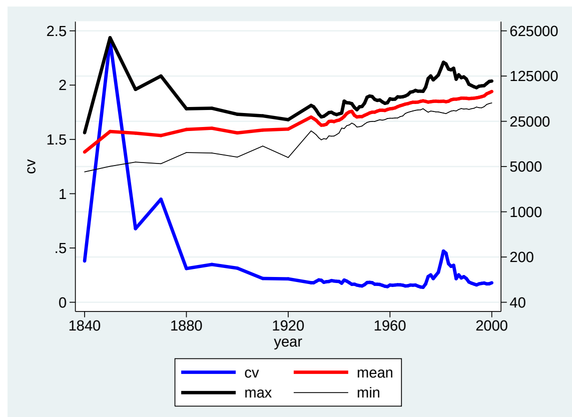
(a) Aggregate



(b) Farm



(c) Manufacturing



(d) NMNF

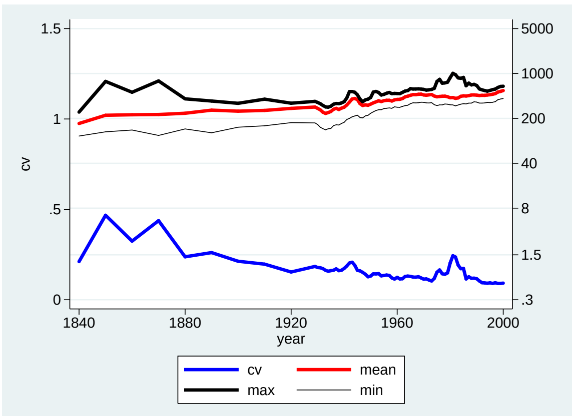
Figure 5: Output Per Worker

3 Growth Accounting Results

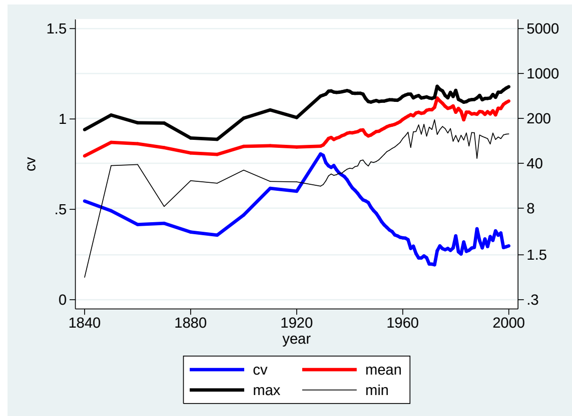
The last ingredient for an accounting exercise is a production function governing the relationship between inputs and outputs. Our goal is to compare our results to the cross-country growth accounting studies, particularly those found in Caselli (2005). Hence we adopt as our baseline specification:

$$y_{it} = A_{it} k_{it}^{\alpha} h_{it}^{1-\alpha} \tag{4}$$

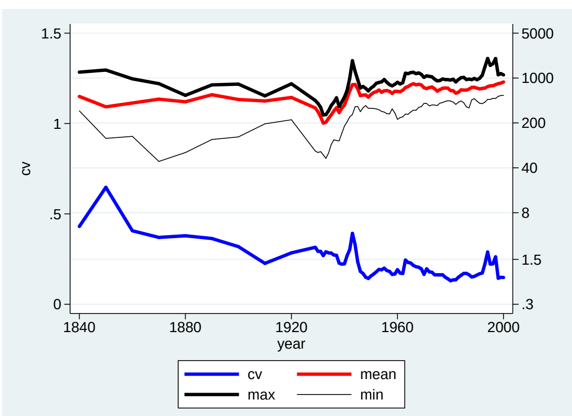
where y, k, and h denote output, capital, and human capital per worker. We take the same $\alpha = 1/3$ as Caselli (2005).[26] Given measured inputs and outputs, A_{it} is constructed as a residual term. It represents the effect of institutions, laws, technology, and other factors that improve the efficiency with which measured inputs can create output.



(a) Aggregate

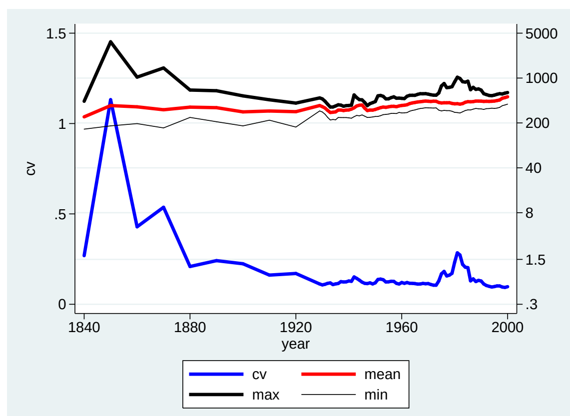


(b) Farm



(c) Manufacturing

S



(d) NMNF

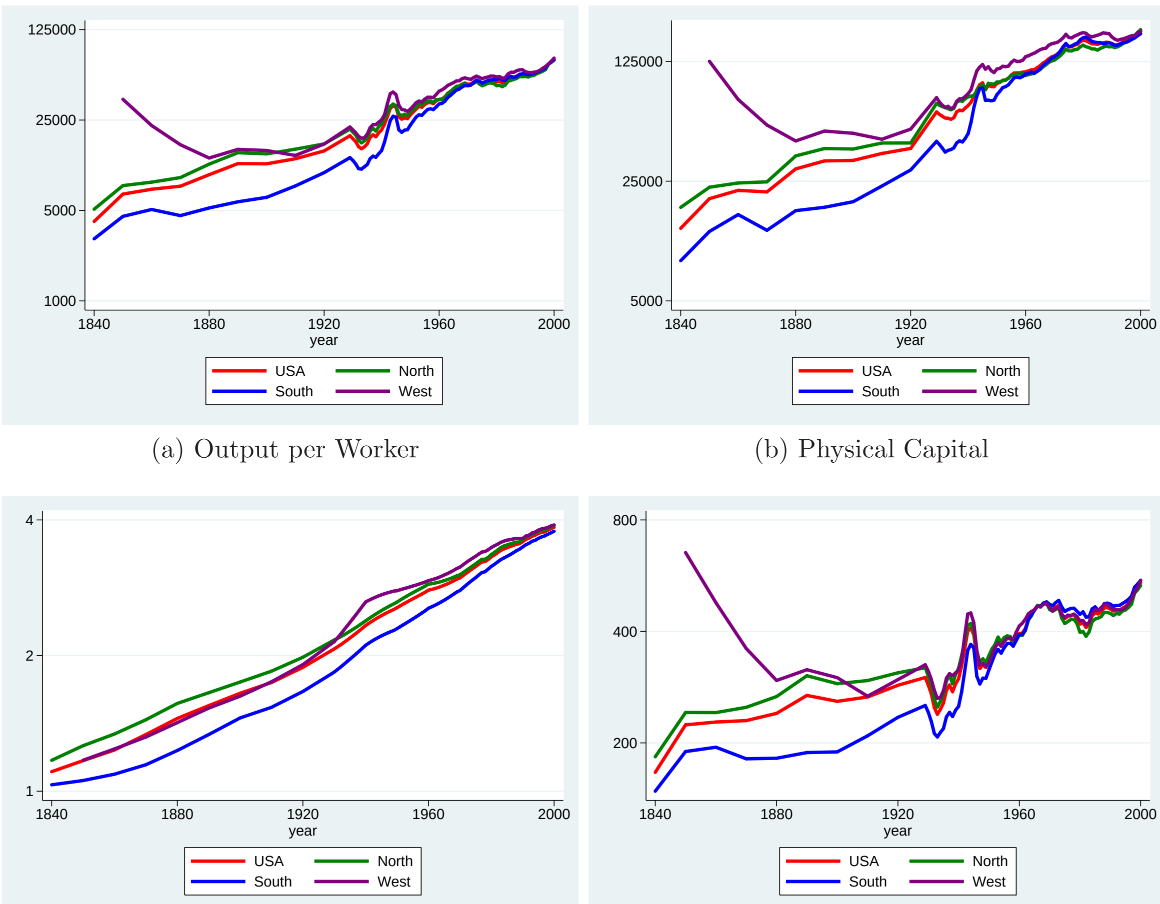
Figure 6: Total Factor Productivity

Figure 6 give the results for TFP. The coefficient of variation decreases nearly uniformly, although there is a brief spike in the late 1970s and early 1980s. The coefficient of variation for the farm sector is much more irregular than for the manufacturing sector or the NMNF sector.

Much of the story of convergence in output per worker and TFP can be told using

[26]Since cross-country studies almost uniformly ignore land, we ignore it here to start. In the following section we see how robust our results are to including land as a valuable input.

three stylized regions of the United States: the North, the South, and the West.[27] Figure 7 contains real output per worker, real physical capital per worker, human capital per worker and TFP by region.



(a) Output per Worker

(b) Physical Capital

(c) Human Capital

(d) TFP

Figure 7: Convergence Across Regions

Briefly, output per worker was high in the 19th century in Western states due to mining activity, which is observed in our measures as high output per worker and levels of TFP not matched since. As the population of Western states boomed and easier mining opportunities diminished, inputs and output per worker decreased in these states towards the level of the North. On the other hand, the South lags in physical capital, human capital, and TFP through most of the period. Output and physical capital decline in 1870 in the aftermath of the Civil War, and converge only slowly. The high variance of output per worker in the

[27]See Section 7 for our definitions of these regions.

nineteenth century are between the Western mining states and the Southern slave states, before and after the Civil War. Our estimates are that TFP had more or less converged to level of the North by 1950, but that physical and human capital continued to lag.[28] By 2000 the Southern worker had 94% of the physical capital of the highest physical capital region, the West, and 97% of the human capital of the highest human capital region, the West. Southern TFP was 99.7% of the highest TFP region, the West, in 2000.

Given our measures of inputs, outputs, and TFP, we now turn to growth accounting exercises that parallel those of the cross-country literature. Between 1840 and 2000, states experience output per worker growth of 1.6% per year. Growth rates tended to be faster in the South and slower in the West than in the North, reflecting convergence of output across states. Our first exercise is a growth accounting exercise, which asks: what proportion of output growth can be accounted for by input growth, and what proportion is left to TFP?

To measure the relative contributions, we use the standard growth accounting equation:

$$g_y = g_a + \alpha g_k + (1-\alpha) g_h = g_a + g_x \tag{5}$$

where g denotes the growth in the corresponding variable. Our measure of the contribution of inputs is simply the fraction of output per worker growth that comes from input growth:

$$\text{Input Share} = \frac{g_x}{g_y}$$

Table 2 gives the results for states, expressed as average annual growth rates. The first column gives the results for the entire 1840-2000 time period. Not all states are available in 1840; for other states, the growth accounting results are calculated from the first year for which we have a complete set of data.[29] Since data in earlier years may be noisier and the set of states with complete records is smaller, the second column gives data for 1900-2000.[30] While only 28 states have complete data for 1840-2000, 47 states do for 1900-2000.[31] By comparison, we have 51 countries with complete data from 1840-2000, and 62 countries with

[28]By 1950 the labor force weighted average TFP levels for the North, South and West are: 353, 318, and 331, respectively. Thus average TFP levels were within 10% of each other. However human capital in the South was only 82% of the West, and 87% of the North in 1950. Physlcal capital per worker in the South was only 61% of the West and 88% of the North in 1950.

[29]See Section 6 for dates of first observation.

[30]The 1900-2000 period is chosen for several reasons: (1) it generally minimizes on worries about initial conditions, (2) it contains the entire 20th century, (3) almost all of the states are available throughout the entire 100 years.

[31]We include Alaska and Hawaii, which we observe from 1950-2000, and Oklahoma, which we observe from 1920-2000. We exclude D.C. from the analysis.

Table 2: Growth Accounting Across States and Countries

	All States, Various Years		Countries, Various Years	
	1840-2000	1900-2000	1840-2000	1900-2000
Unweighted				
g_y	1.61%	1.90%	1.05%	1.14%
g_k	1.42%	1.70%	1.64%	1.77%
g_h	0.80%	0.88%	0.78%	0.85%
g_x	1.01%	1.15%	1.07%	1.17%
g_a	0.60%	0.75%	-0.03%	-0.03%
Input share	.63	.61	1.02	1.02
Labor Force Weighted				
g_y	1.54%	1.92%	0.99%	1.32%
g_k	1.35%	1.75%	1.10%	1.44%
g_h	0.80%	0.87%	0.55%	0.72%
g_x	0.98%	1.16%	0.75%	0.98%
g_a	0.56%	0.75%	0.24%	0.33%
Input share	.64	.61	.76	.75
Obs	50	50	168	168

Note: y denotes output per worker, $x = k^{\alpha}h^{1-\alpha}$ denotes weighted inputs per worker, and a denotes TFP. For cross country comparisons we use data from Tamura, Dwyer, Devereux, and Baier (2011), which essentially is Maddison data on output, and authors' construction of data for physical capital and human capital based on a variety of recognized data sources.

complete data from 1900-2000. Growth was slightly faster for the second sample, but the key results are very similar for the two samples. A bit more than three-fifths of output per worker growth is accounted for by input growth; capital growth was slightly more important for input growth. This is true for both the unweighted results and the weighted results.[32]

To put these figures in context, we compare them to similar accounting results for countries in the last two columns. We take the data in Tamura, Dwyer, Devereux, and Baier (2011), hereafter TDDB, and construct two samples for a growth accounting exercises. In the first we use all years of data for 168 countries, with final output year of 2000, and for the second sample we restrict to years 1900-2000.[33]

We find that the average country grew slower than the average state, about 1.0% per year. For countries, slightly more of growth was accounted for by input growth, 75%-

[32] In the weighted cases, we weight by labor force in the terminal year of observation.

[33] We include all countries contained in TDDB, but we start in 1840 (1900), or the first year of observation, whichever is later. In TDDB the final year of observation is 2007. However in order to make the samples more comparable, we restricted the last year of data to 2000.

102%.[34] As with the states, input growth across countries was more driven by capital stock growth than human capital growth. Overall, we take away that growth accounting across states is similar to growth accounting across countries, with state inputs accounting for around three-fifths and country inputs accounting for three-fourths of total growth.

3.1 Growth Variance Decomposition

Growth in inputs accounts for most of growth in output in states and countries. However, Klenow and Rodriguez-Clare (1997) and Easterly and Levine (2002) argue that most of the variation of growth rates across countries is accounted for by variation in TFP growth rates rather than variation in input growth rates. Here, we pursue this variation of growth accounting, which asks: what proportion of the variation in growth rates can be accounted for by the variation in input growth rates, and what proportion is left to TFP?

We consider four measures of the role of inputs. First, the ratio of the variances:

$$\text{Input Share} = \frac{\text{var}[g_x]}{\text{var}[g_y]}$$

and second, a measure which assigns half of the covariance term to inputs:

$$\text{Input Share} = \frac{\text{var}[g_x] + \text{covar}[g_x, g_a]}{\text{var}[g_y]}$$

Third we compute the 90-10 ratio for output growth rates with the 90-10 ratio for input growth rates. While it is the case that the cross-country comparison is impossible, because there are significant numbers of negative output growth rates, but all inputs grow, it is useful at the state level. Fourth we use a different assignment method for the covariance term. We follow Baier, Dwyer and Tamura (2006) and TDDB (2011) in allowing the data to inform us as to the assignment of the covariance term.

The correlation of growth rates of inputs and total factor productivity growth is not 0. There are two sets of theories that imply that the correlation between input growth and TFP growth is caused by one or the other. For example, neoclassical growth models with exogenous technological progress, and Romer (1990), with endogenous technological progress, imply that factor accumulation is induced by the growth in TFP. On the opposite end of the theoretical divide, Romer (1986), Lucas (1988), and Tamura (2002, 2006) con-

[34]We generally prefer the labor force weighted results compared to the unweighted results. There are many more small countries in Africa, Latin America and Asia that have relatively shorter horizons than the principal large population countries.

struct theories that show that physical capital accumulation or human capital accumulation produces endogenous TFP growth. Thus these sets of theories imply that the correlation between TFP growth and input growth are due to input growth and hence the correlated or predictable component should be assigned to input growth.

Under the view that TFP growth induces factor accumulation, and that the predictable or correlated portion of input growth should be assigned to TFP growth, the share of growth of output per worker can be written as:

$$1 = \frac{(sd(g_a) + sd(g_x)\rho_{g_x,g_a})^2}{var(g_y)} + \frac{(1-\rho^2_{g_x,g_a})var(g_x)}{var(g_y)} \tag{6}$$

where the first term is now a plausible upper bound on the proportion of the variation in growth rates of output per worker caused by variation in growth rates of TFP.[35] At the other end of the theoretical spectrum, the predictable or correlated component of TFP growth arises from endogenous factor accumulation. Assigning this predictable component to factor accumulation produces the following variance decomposition:

$$1 = \frac{(sd(g_x) + sd(g_a)\rho_{g_x,g_a})^2}{var(g_y)} + \frac{(1-\rho^2_{g_x,g_a})var(g_a)}{var(g_y)} \tag{7}$$

The first term is now the proportion of the variation of growth rates of output per worker that explained by variation in input growth.[36] We produce both plausible upper bounds, and compute an average of input share. Thus we produce for input growth:

$$\text{BDT Average Decomposition} = \frac{(sd(g_x) + sd(g_a)\rho_{g_x,g_a})^2}{2var(g_y)} + \frac{(1-\rho^2_{g_x,g_a})var(g_x)}{2var(g_y)} \tag{8}$$

Our results are given in Table 3. As for growth accounting, we study two eras for states, 1840-2000 and 1900-2000.[37] The results for states are similar across the two time periods. If we average all four measures of input share we find that variation in input growth rates account for 40% (50%) of output per worker growth rate variation in the 1840-2000 (1900-2000) period. Thus between 50% and 60% of the variation in output per worker growth rates is accounted for by variation in TFP growth rates. We also present results for the two

[35] One way of seeing that the least squares decomposition holds for this representation is to note that the variance decomposition is $var(y) = \beta^2_{y,a}var(a) + var(e_{y|a})$, where $\beta_{y,a}$ is the regression coefficient from a regression of y on a and $e_{y|a}$ is the regression residual.

[36] One way of seeing that the least squares decomposition holds for this representation is to note that the variance decomposition is $var(y) = \beta^2_{y,x}var(x) + var(e_{y|x})$, where $\beta_{y,x}$ is the regression coefficient from a regression of y on x and $e_{y|x}$ is the regression residual.

[37] As with the growth accounting, the initial year is either 1840 (1900) or first year of observation.

samples of countries constructed in the previous subsection. For countries, input growth variation accounts for 36% of the variation of growth rates of output per worker. This is true for both the 1840-2000 and the 1900-2000 period. The rise in the input share for the states from 40% to 50% is perhaps not surprising. The end of slavery and the settlement of the entire continental US by 1900 makes the second period one of more homogeneous state institutions and sectoral composition.[38]

Table 3: Variation of Growth Accounting Across States and Countries

	States		Countries	
Year	1840-2000	1900-2000	1840-2000	1900-2000
Relative Variance				
σ_y^2	0.0020%	0.0016%	0.0225%	0.0238%
σ_x^2	0.0002%	0.0003%	0.0058%	0.0059%
Input share	.12	.21	.26	.25
Relative Variance, Accounting for Covariance				
σ_y^2	0.0020%	0.0016%	0.0225%	0.0238%
$\sigma_x^2+\sigma_{x,a}$	0.0006%	0.0007%	0.0091%	0.0097%
Input share	.32	.42	.41	.41
Comparison of 90/10 Ratios				
g_y^{90}/g_y^{10}	2.10	1.63	-	-
g_x^{90}/g_x^{10}	1.51	1.48	-	-
Input share	.72	.91	-	-
BDT Average Decomposition				
Input share	.44	.46	.43	.43
Avg. input share	.40	.50	.36	.36
Avg. input share without 90/10 ratio	.30	.36	.36	.36
Obs	50	50	168	168

Overall, the cross-state results are similar to the cross-country results, with inputs having slightly more ability to explain cross state growth rate variation than cross-country growth rate variation. For the states, about three fifths of growth in output per worker is accounted for by input growth, and something like 40 percent of the cross sectional variation in growth rates is explained by cross sectional variation in input growth rates. By contrast, in the cross-country data, over three-fourths of output per worker growth is accounted for by input per worker growth, and almost 40 percent of the variation in growth of output per

[38]Farming becomes less important between 1840 and 1900. In 1840 62.5% of the workforce is in farming in the states, and by 1900 farming had declined to 38%. Farming, with its non tradable inputs of land and climate, would behave differently than manufacturing, with fewer locational specific production amenities.

worker can be accounted for by variation in input per worker growth rates.

4 Robustness Checks: Regions, Sub-Periods, Alternative Inputs, and Sectors

We check the robustness of our results by considering regions, time sub-periods, alternative inputs, and sectors. In general we find that the aggregate results are robust to all of these checks. We present the results in this section in a slightly unusual manner. First we present the regional analysis, growth accounting and the variance decomposition of growth. Second we present growth accounting and variance decomposition of growth for time sub-periods. Next we introduce land as an input in farming and produce an alternative aggregate input. We then conduct growth accounting and variance decomposition of growth with this new input. We summarize these results as well as sectoral results in the next sub-section. We provide this somewhat early summary because of the overwhelming number of statistics can be numbing. Finally we present the detailed sectoral growth accounting and variance decomposition of sectoral growth.

4.1 Regions

We examine the state growth rates by regions, North, South and West. Again we examine growth from 1840 (or initial year) to 2000 and from 1900 (or initial year) to 2000. These results are contained in Table 4. As evident from Figure 7, the South and the North grow faster than the West. Input growth accounts for between 57% and 69% of growth in the unweighted case, and between 57% and 93% of growth in the weighted case. Generally inputs explain the highest proportion of growth in the slowest growing region, West, and the smallest proportion of growth in the fastest growing region, South. Despite this, in no region or time period do inputs explain less than 57% of average growth. As before physical capital growth is more rapid than human capital growth.

Table 5 contains the variance decomposition of growth rates by regions of the US. We present the results by region for the two different time periods, 1840-2000 and 1900-2000.[39] Generally the 1840-2000 period has more variable growth rates than the 1900-2000 period. This greater variance is more difficult to capture based on variation in input growth rates. For the 1840-2000 period, the average shares across the regions of growth rate variations "explained" by input growth rate variations are 14% (relative variance),

[39] Or from first year of data availability to 2000.

Table 4: Growth Accounting Across US Regions

	North		South		West	
	1840-2000	1900-2000	1840-2000	1900-2000	1840-2000	1900-2000
Unweighted						
g_y	1.57%	1.66%	1.91%	2.39%	1.31%	1.67%
g_k	1.39%	1.56%	1.77%	2.22%	1.04%	1.28%
g_h	0.78%	0.82%	0.81%	0.94%	0.83%	0.90%
g_x	0.98%	1.06%	1.13%	1.36%	0.90%	1.02%
g_a	0.59%	0.60%	0.78%	1.03%	0.41%	0.64%
Input share	.63	.64	.59	.57	.69	.61
Labor Force Weighted						
g_y	1.63%	1.67%	1.93%	2.41%	0.77%	1.62%
g_k	1.43%	1.58%	1.79%	2.23%	0.54%	1.34%
g_h	0.77%	0.80%	0.82%	0.96%	0.80%	0.88%
g_x	0.99%	1.06%	1.14%	1.38%	0.72%	1.04%
g_a	0.64%	0.62%	0.79%	1.03%	0.05%	0.59%
Input share	.61	.63	.59	.57	.93	.64
Obs	21	21	16	16	13	13

30% (with covariance), 70% (with 90-10 ratios), and 38% (BDT average decomposition). In contrast the second period has higher average shares: 35% (relative variance), 35% (with covariance), 99% (90-10 ratios), and 36% (BDT average decomposition). Taking a grand average across all regions of each of the four measures for both time periods reveals that input growth variance captures 38% (1840-2000) and 51% (1900-2000) of growth rate variance. Without the 90-10 ratios, input growth variation captures 27% (1840-2000) and 35% (1900-2000) of output growth variation. Thus between 49% to 73% of output growth variation is explained by TFP growth rate variation. Overall something like 40% of output growth rate variations are explained by input growth rate variations, and 60% are explained by TFP growth rate variation.[40]

4.2 Sub-Periods

We examine the time series of growth from 1840-2000 in 40 year increments. Table 6 contains the growth results for the states as well as for countries in successive 40 year subperiods.[41] We compare these results to those obtained from the cross country data of

[40]This result is similar to the 50% share of in BDT average decomposition found in TDDB.

[41]We examine the annualized growth rates over 1840-1880, 1880-1920, 1920-1960, 1960-2000.

Table 5: Variation of Growth Accounting Within Regions

	North		South		West	
Year	1840-2000	1900-2000	1840-2000	1900-2000	1840-2000	1900-2000
Relative Variance						
σ_y^2	0.0005%	0.0001%	0.0005%	0.0009%	0.0045%	0.0006%
σ_x^2	0.0001%	0.0001%	0.0001%	0.0002%	0.0004%	0.0001%
Input share	.11	.61	.20	.25	.10	.18
Relative Variance, Accounting for Covariance						
σ_y^2	0.0005%	0.0001%	0.0005%	0.0009%	0.0045%	0.0006%
σ_x^2+$\sigma_{x,a}$	0.0001%	0.0000%	0.0002%	0.0004%	0.0014%	0.0001%
Input share	.17	.34	.42	.45	.30	.25
Comparison of 90/10 Ratios						
g_y^{90}/g_y^{10}	1.44	1.14	1.36	1.43	5.43	1.34
g_x^{90}/g_x^{10}	1.18	1.18	1.27	1.33	1.91	1.32
Input share	.82	1.04	.93	.93	.35	.99
BDT Average Decomposition						
Input share	.19	.36	.48	.48	.47	.25
Avg. input share	.32	.59	.51	.53	.31	.42
Avg. input share without 90/10 ratio	.16	.44	.37	.39	.29	.23
Obs	21	21	16	16	13	13

TDDB. In both cases growth accelerates from the 1840-80 period to the 1920-1960 period. For the US states, growth declines from 2.42% in the 1920-1960 period to 1.76% in the 1960-2000 period, and a similar decline occurs in the labor force weighted state case. However for the world, the weighted case shows acceleration in economic growth from 1.11% to 2.05%, whereas the unweighted case shows decline from 1.97% to 1.08%.[42] Input growth accounts for between 54% and 80% of growth in the US states. Similarly for countries, growth of inputs is responsible for between 45% and more than 100% of growth in the unweighted case, and between 41% and 77% in the weighted case. The average shares of growth from inputs for states and countries are: 68% (61% weighted) and 73% (60% weighted), respectively.

Table 7 contains the variance of growth decomposition for states and countries. The variation in growth rates among the states of the US over the different subperiods is similar to the experience across countries. Ignoring the 90/10 ratios, the average input shares in relative variance, relative variance with covariance, and BDT average variance decompo-

[42]This is driven by the decline in many former Soviet republics, first observed in 1970, and in many Sub-Saharan African countries, and the acceleration of growth in China and India.

sition for the US states are 20%, 38% and 44%, respectively. For the cross country data, the average input shares are: 47%, 42% and 41%, respectively. Averaging the states across these three shares produces 34%. For the cross-country data, for relative variance, relative variance with covariance and BDT average variance decomposition, the input shares are: 26%, 37% and 31%, respectively. Averaging produces input shares of 34% for the US state data and 31% for the cross country data. Thus inputs are slightly better at explaining output growth variation for the US states than in the cross-country data.

Perhaps it is surprising again that there is not a larger discrepancy in favor of the US states compared with the cross-country data. If we restrict the time period even more to only the 1880-2000 years we find that the average input shares for the US states are: 23%, 40% and 43% respectively. For the cross-country data the average input share are: 27%, 36% and 37%, respectively. Averaging across these three measures produces: 35% (US) and 33% (cross country).[43] A priori we expected the greater mobility of capital and labor, more common institutions, common language, common currency and common trade policy for the US states to have a larger effect on the share of growth rate variation explained by input growth variation. Variation in TFP growth rates are more important than input growth rate variation in explaining variation in output per worker growth rates for the US states.

4.3 Alternative Input per Worker: Land

In this section we incorporate our calculations on land per farm worker. We compute input per worker by first producing input per worker by sector, and then weighting by the sectoral labor force. Thus we produce:

$$x_{farm} = \mathcal{L}^{.19} k_{farm}^{.21} h_{farm}^{.60} \tag{9}$$

$$x_{man} = k_{man}^{.33} h_{man}^{.67} \tag{10}$$

$$x_{nmnf} = k_{nmnf}^{.33} h_{nmnf}^{.67} \tag{11}$$

$$x = x_{farm} * s_{farm} + x_{man} * s_{man} + x_{nmnf} * (1 - s_{farm} - s_{man}) \tag{12}$$

where s_{farm} is the share of the workforce working in farming, s_{man} is the share of the workforce working in manufacturing and the last term is the share of workforce in NMNF.[44] We present the growth accounting results in the aggregate for 1840-2000 and 1900-2000 in

[43] The average for the 90-10 ratios is 95% for the US states.

[44] In the farm sector we use the factor shares used in Caselli and Coleman (2001) and Jorgenson and Gollop (1992).

Table 6: Growth Accounting Across States and Countries

	All States, Various Years				Countries, Various Years			
	1840-80	1880-1920	1920-60	1960-2000	1840-80	1880-1920	1920-60	1960-2000
Unweighted								
g_y	0.96%	1.07%	2.42%	1.76%	0.73%	1.05%	1.97%	1.08%
g_k	1.07%	0.71%	2.67%	1.22%	0.84%	1.07%	1.54%	1.97%
g_h	0.63%	0.68%	1.01%	0.82%	0.26%	0.29%	0.57%	1.09%
g_x	0.78%	0.69%	1.56%	0.95%	0.45%	0.56%	0.99%	1.37%
g_a	0.19%	0.38%	0.86%	0.81%	0.28%	0.49%	0.99%	-0.28%
$\frac{g_x}{g_y}$	80%	65%	64%	54%	62%	54%	50%	126%
Weighted by labor force								
g_y	1.67%	1.05%	2.31%	1.77%	0.30%	0.68%	1.11%	2.05%
g_k	1.54%	0.63%	2.58%	1.29%	0.35%	0.42%	0.72%	2.59%
g_h	0.68%	0.65%	0.99%	0.82%	0.13%	0.21%	0.58%	1.09%
g_x	0.96%	0.64%	1.52%	0.97%	0.21%	0.29%	0.67%	1.59%
g_a	0.71%	0.41%	0.80%	0.79%	0.09%	0.40%	0.44%	0.46%
$\frac{g_x}{g_y}$	58%	61%	66%	55%	70%	42%	61%	77%
Obs	35	45	48	50	53	60	68	165

Table 8. We repeat the same growth accounting results from the cross-country data for ease of comparison. Overall we find that the inclusion of land as an input has almost no effect on the results. The share of growth accounted for by input growth is 60% compared with 63% before. In Table 9 we report the variance decomposition results. With the exception of a dramatic decline in power to explain growth rate variation over the 1840-2000 period, there is no change in the variance decomposition results. Averaging the four input shares yields 28% (1840-2000) and 50% (1900-2000) compared with the earlier estimates of 40% (1840-2000) and 50% (1900-2000). Thus the inclusion of land as input in aggregate output does not help to explain variation in output growth rates as long as the maintained assumption of common factor shares in farming are assumed.

4.4 Quick Summary

Table 10 contains an overview of our robustness results. It contains the average input share explaining growth rates and the variation in growth rates. The top half of the table summarizes the results for growth rates. It includes results with and without weights, with or without land, by region of the country, by time period and by sector. The bottom half

Table 7: Variation of Growth Accounting Across States & Countries

	All States, Various Years				Countries, Various Years			
Year	1840-80	1880-1920	1920-60	1960-2000	1840-80	1880-1920	1920-60	1960-2000
Relative Variance								
σ_y^2	0.0997%	0.0032%	0.0035%	0.0015%	0.0150%	0.0049%	0.0341%	0.0357%
σ_x^2	0.0104%	0.0007%	0.0006%	0.0005%	0.0035%	0.0018%	0.0090%	0.0071%
Input share	.10	.22	.17	.31	.24	.36	.26	.20
Relative Variance, Accounting for Covariance								
σ_y^2	0.0997%	0.0032%	0.0035%	0.0015%	0.0150%	0.0049%	0.0341%	0.0357%
$\sigma_x^2 + \sigma_{x,a}$	0.0314%	0.0012%	0.0012%	0.0007%	0.0016%	0.0015%	0.0124%	0.0138%
Input share	.31	.36	.34	.49	.11	.32	.36	.39
Comparison of 90/10 Ratios								
g_y^{90}/g_y^{10}	-	3.19	1.90	1.70	-	3.79	4.72	-
g_x^{90}/g_x^{10}	-	3.09	1.43	1.91	-	46.3	4.36	-
Input share	-	.97	.75	1.12	-	12.22	0.92	-
BDT Average Decomposition								
Input share	.47	.39	.40	.49	.13	.32	.37	.43
$\overline{share}$	.30	.49	.41	.60	.16	3.30	.48	.34
$\overline{share}_{\text{no 90-10}}$	.30	.32	.30	.43	.16	.33	.33	.34
Obs	35	45	48	50	53	60	68	165

of the table summarizes the results for explaining growth rate variations, it includes results with and without the 90/10 ratios, with or without land, by region of the country, by time period and sector. It is useful to summarize the large amount of results now, to introduce some perspective. For growth rates, the inputs explain about three fifths of growth in output per worker; this is robust to time periods, 1840-2000 and 1900-2000, as well as sub-periods, consecutive forty year periods beginning with 1840, and regions of the country. It is also robust whether one looks at weighted or unweighted data. For sectoral growth, inputs explain about two thirds of growth for manufacturing and NMNF, and about four ninths of growth in farming. Again these results are robust to time periods, sub-periods, weighted or unweighted, and regions. For variation of growth rates, the inputs explain about forty percent of the variation. There is more variation in the share than in the growth case, but

Table 8: Growth Accounting Across States (with Land) and Countries

	All States, Various Years		Countries, Various Years	
	1840-2000	1900-2000	1840-2000	1900-2000
Unweighted				
g_y	1.61%	1.90%	1.05%	1.14%
g_k	1.42%	1.70%	1.64%	1.77%
g_h	0.80%	0.88%	0.78%	0.85%
$g_{\mathcal{L}}$	1.16%	1.39%		
g_x	0.97%	1.09%	1.07%	1.17%
g_a	0.64%	0.81%	-0.03%	-0.03%
Input share	.60	.57	1.02	1.02
Labor Force Weighted				
g_y	1.54%	1.92%	0.99%	1.32%
g_k	1.35%	1.75%	1.10%	1.44%
g_h	0.80%	0.87%	0.55%	0.72%
$g_{\mathcal{L}}$	0.74%	1.28%		
g_x	1.00%	1.11%	0.75%	0.98%
g_a	0.54%	0.81%	0.24%	0.33%
Input share	.65	.58	.76	.75
Obs	50	50	168	168

nonetheless this result is robust to time period, sub-period, sectors and with or without the 90/10 ratio.

4.5 Sectoral Growth

To check the robustness of our results on real output per worker growth, we use the three sectors, farming, manufacturing and NMNF, as alternative output measures to examine.[45] We use a different technology for farming, including land as an input, compared with manufacturing and NMNF. We allow TFP in each of the three sectors to grow at different rates, and see if input growth and the variation in input growth captures sectoral output growth and variations in sectoral output growth similar to our aggregate output per worker growth. We use the differential importance of land, as well as differences in human capital and physical capital by sector to examine growth in each of the three sectors, farming, manufacturing and NMNF. We first examine the growth record between the two time periods, 1840-2000 and 1900-2000. We next examine the sectoral growth by regions as well

[45] Schoellman and Tamura (2011) examine development accounting by sector.

Table 9: Variation of Growth Accounting Across States (with Land) and Countries

	States		Countries	
Year	1840-2000	1900-2000	1840-2000	1900-2000
Relative Variance				
σ_y^2	0.0020%	0.0016%	0.0225%	0.0238%
σ_x^2	0.0003%	0.0004%	0.0058%	0.0059%
Input share	.15	.22	.26	.25
Relative Variance, Accounting for Covariance				
σ_y^2	0.0020%	0.0016%	0.0225%	0.0238%
$\sigma_x^2+\sigma_{x,a}$	0.0002%	0.0006%	0.0091%	0.0097%
Input share	.09	.39	.41	.41
Comparison of 90/10 Ratios				
g_y^{90}/g_y^{10}	2.10	1.63	-	-
g_x^{90}/g_x^{10}	1.56	1.55	-	-
Input share	.74	.95	-	-
BDT Average Decomposition				
Input share	.10	.43	.43	.43
Avg. input share	.27	.50	.36	.36
Avg. input share without 90-10 ratio	.11	.35	.36	.36
Obs	50	50	168	168

as time periods. Finally we examine the sectoral growth in 40 year periods.

Table 11 contains the results of growth accounting by sectors. There is evidence that output per worker growth accelerated between the 19th and 20th centuries. In all three sectors, output per worker growth is more rapid in the 1900-2000 period than in the longer 1840-2000 period. Some of the acceleration is due to the high initial values of output per worker in the mining states, but this only affects the NMNF sector. Manufacturing and farming both exhibit acceleration in output per worker growth.

Perhaps surprising to some is the rapid growth rate of farming output per worker in both the long period, 1840-2000, and the short period, 1900-2000. This is true whether the data are weighted or not. However further reflection indicates that the exit out of farming into manufacturing and services was made possible by the rapid increase in output per worker in farming. Perhaps also surprising to some is that this sector experienced the most rapid advance in TFP, with growth rates equal to or exceeding the growth rate of NMNF output per worker![46] Again this rapid expansion in TFP is what made possible the exit of

[46]We are including land as an input in farming throughout all of the growth accounting and variance decomposition of growth rates in farming.

workers from farming into manufacturing and NMNF. Inputs explain about 45% of farm output growth.

There is a great deal of similarity between manufacturing and NMNF. While labor force weighted manufacturing output per worker grows at a more rapid rate than NMNF output per worker, about 1.9% versus 1.5%, the share of growth captured by input growth is quite similar. Input growth explains 70% and 75% of average output growth in manufacturing and NMNF, respectively.

Table 12 reports the results from the decomposition of sectoral growth rate variations. In the 1840-2000 period, the grand average of all four measures of variation in output growth rates explained by variation in input growth rates produces 54%, 34% and 28% in farming, manufacturing and NMNF, respectively. There is a slight increase in the ability to account for growth rate variations in the 1900-2000 period. The input shares are 57%, 39% and 46% in farming, manufacturing and NMNF, respectively. Generally speaking variation in growth rates of inputs are able to explain something on the order of 47% of the variation in output growth rates over the last century, and about 39% of output growth rate variation from 1840-2000. Both the growth accounting results and the decomposition of output growth variation from the sectors are consistent with the aggregate results contained herein on the states, as well as those in cross-country analysis, c.f. Baier, Dwyer and Tamura (2006), Tamura, Dwyer, Devereux, and Baier (2011).

4.6 Region & Sector

We repeat the sectoral analysis, but for the three regions, North, South and West. In both the unweighted and weighted cases, there is evidence of convergence in levels, see Table 13. The South is the poorest region of the three, and it has the most rapid output growth in the highest average output per worker sector, NMNF. Ignoring farming, input growth in both manufacturing and NMNF explains between 62% and 104% of average growth in these two sectors.[47] Farming, by contrast, has an input share of growth of 46%. The results are fairly similar in the 1900-2000 period. Share of growth accounted for by input growth in manufacturing and NMNF are very similar, 66%, and farming is significantly lower 44%.

Table 14 contains the sectoral growth variance decomposition by region and by time period. The highest share of growth rate variation explained by variation in input growth occurs in farming. This holds whether we look at the entire period, 1840-2000, or the past century, 1900-2000, and it holds whether we include the 90-10 ratio or not. For farming,

[47]This ignores the weighted NMNF sector results. These are driven by the negative growth in NMNF for California.

averaging over the three regions, we find that inputs can explain 31% (54% including the 90-10 ratios) of output growth variation in the 1840-2000 period, and 48% (66% including the 90-10 ratios) of output growth variation in the 1900-2000 period. The comparable numbers for NMNF are: 30% (35% including the 90-10 ratios) over the 1840-2000 period and 27% (42% including the 90-10 ratios) over the 1900-2000 period. The numbers for manufacturing are: 18% (30% including the 90-10 ratios) over the 1840-2000 period and 16% (32% including the 90-10 ratios) over the 1900-2000 period. Thus input growth variation are capable of explaining something like 30% (44% including 90-10 ratios) of output growth variations. These results are comparable to those found in Baier, Dwyer and Tamura (2006), Klenow and Rodriguez-Clare (1997), Weil (2009), and somewhat less than those in Tamura, Dwyer, Devereux, and Baier (2011).

Finally we report on growth by sector in the four forty year periods, 1840-80, 1880-1920, 1920-60 and 1960-2000. We examine the behavior with all the states available instead of regionally for purposes of maintaining sample size. Table 15 contains the growth accounting results.[48] For farming, input growth accounts for between 33% and 138% (29% and 100% in the weighted case) of output per worker growth. Manufacturing has negative average growth in the first period, thus ignoring this period input growth explains 61% (in the unweighted case) and 64% (in the weighted case) of average manufacturing growth. In the NMNF sector, there is negative growth in the 1880-1920 period, and ignoring that period inputs explain 72% of output per worker growth in the unweighted case and 63% of per worker output growth in the weighted case.

Table 16 contains the variance decomposition results. We see that manufacturing and NMNF sectors look similar. Averaging over all four measures we find that variations in input growth rates explain 48% and 64% of the variation in output per worker growth rates in manufacturing and NMNF, respectively. Ignoring the 90-10 ratio, the remaining three measures show that variations in input growth rates explain 40% and 33% of output growth rate variations in manufacturing and NMNF, respectively. The overall input share in farming is 16% without the 90-10 ratio, and 81% with the 90-10 ratio. Certainly farming has more specific factors than either manufacturing or NMNF. Land quality, climate and weather are local variables, whereas assembly of autos, generating electricity using coal or providing banking services are likely less location specific. While we are controlling for land quality, we are not able to control for weather, climate, the ease of clearing and improving land, the ease of transportation of perishable items. Perhaps these are features that make

[48] In the weighted case, we use the labor force in the terminal year of the observation, e.g. 2000, 1960, 1920, etc.

farming more difficult to explain than manufacturing and NMNF.

5 Conclusion

Growth accounting has offered improvements to our understanding of cross-country output per worker differences. This paper has used the methods of cross-country growth accounting exercises in an alternative laboratory, the states of the United States, which offer an interesting alternative environment. While output per worker differences across states in the 19th century were large, output per worker converged in the 20th century in a way not seen across countries. The Censuses of Manufactures and Agriculture offer a long history of direct measurement of inputs and outputs at the sectoral level suitable for sectoral accounting exercises. Finally, the factors not accounted for - institutional and legal differences, slavery, and so on - are easier to identify across states than across countries.

Of all the conclusions derived from the growth accounting literature, perhaps the most prominent is the primacy of TFP. Using conventional estimates of factor shares, there is insufficient variation in the growth rate of inputs to account for even half of the diversity in the level or growth rate of output per worker across countries.[49] Indeed, typical figures put the role for input variation closer to one-quarter to one-third of the total. This paper offers additional confirmation to that conclusion. Included in this robustness check is the addition of land as a valuable input of production in farming. Whether considering the augmented aggregate input or in farming alone, we find surprisingly small additional information contained in land. We reiterate the belief that this points towards state specific-crop specific production technologies, that are beyond the scope of this paper. Variation in input growth accounts for between one third to four ninths of output per worker variation. Hence most of the variation in output per worker growth rates is accounted for through variation in TFP growth rates. The primacy, albeit reduced, of TFP holds for different regions, different sectors and different periods of U.S. history.

[49]Tamura, Dwyer, Devereux, and Baier (2011) is an exception. With a longer time series over 168 countries, they find that about 50% of the growth rate variations are explained by input growth rate variations.

Table 10: Average Input Share of Growth and Growth Variation

	Period	Category	US no land	US land	North	South	West	1840-80	1880-1920	1920-60	1960-2000
Growth Accounting: Average Input Share											
Unweighted	1840-2000	All	0.63	0.60	0.63	0.59	0.69	0.80	0.65	0.64	0.54
		Farming		0.48	0.44	0.47	0.57	1.38	0.76	0.33	0.48
		Manufacturing	0.70		0.74	0.74	0.63	-0.64	0.63	0.60	0.60
		NMNF	0.79		0.77	0.70	1.04	0.96	-634.19	0.71	0.49
	1900-2000	All	0.61	0.57	0.64	0.57	0.61				
		Farming		0.46	0.44	0.46	0.50				
		Manufacturing	0.66		0.68	0.67	0.62				
		NMNF	0.65		0.65	0.62	0.72				
Weighted	1840-2000	All	0.64	0.65	0.61	0.59	0.93	0.58	0.61	0.66	0.55
		Farming		0.42	0.40	0.46	0.41	0.52	1.00	0.29	0.48
		Manufacturing	0.76		0.81	0.73	0.74	-10.36	0.68	0.63	0.61
		NMNF	0.90		0.75	0.71	-0.85	0.68	-8.59	0.70	0.50
	1900-2000	All	0.61	0.58	0.63	0.57	0.64				
		Farming		0.42	0.40	0.46	0.41				
		Manufacturing	0.68		0.71	0.68	0.62				
		NMNF	0.65		0.64	0.62	0.73				
Variation of Growth: Average Input Share											
With 90/10 ratio	1840-2000	All	0.40	0.27	0.32	0.51	0.31	0.30	0.49	0.41	0.60
		Farming		0.54	0.43	0.55	0.64	1.95	0.79	0.19	0.32
		Manufacturing	0.34		0.29	0.29	0.32	0.27	0.65	0.37	0.65
		NMNF	0.28		0.37	0.41	0.27	1.01	0.33	0.38	0.55
	1900-2000	All	0.50	0.50	0.59	0.53	0.42				
		Farming		0.57	0.48	0.79	0.70				
		Manufacturing	0.39		0.37	0.24	0.35				
		NMNF	0.46		0.47	0.47	0.32				
Without 90/10 ratio	1840-2000	All	0.30	0.11	0.16	0.37	0.29	0.30	0.32	0.30	0.43
		Farming		0.31	0.10	0.41	0.43	0.10	0.19	0.04	0.31
		Manufacturing	0.24		0.15	0.19	0.20	0.27	0.62	0.26	0.47
		NMNF	0.29		0.29	0.33	0.27	0.30	0.33	0.31	0.37
	1900-2000	All	0.36	0.35	0.44	0.39	0.23				
		Farming		0.36	0.24	0.64	0.55				
		Manufacturing	0.27		0.22	0.05	0.22				
		NMNF	0.33		0.32	0.35	0.15				

Table 11: Growth Accounting By Sectors

	Farming		Manufacturing		NMNF	
	1840-2000	1900-2000	1840-2000	1900-2000	1840-2000	1900-2000
Unweighted						
g_y	2.28%	2.76%	1.80%	2.02%	1.01%	1.45%
g_k	2.09%	2.53%	2.29%	2.37%	0.79%	1.15%
g_h	0.74%	0.80%	0.76%	0.82%	0.80%	0.84%
$g_{\mathcal{L}}$	1.16%	1.37%				
g_x	1.10%	1.27%	1.27%	1.33%	0.79%	0.94%
g_a	1.18%	1.49%	0.53%	0.69%	0.22%	0.50%
Input share	.48	.46	.70	.66	.79	.65
Labor Force Weighted						
g_y	2.16%	2.84%	1.75%	1.90%	0.83%	1.49%
g_k	1.75%	2.45%	2.50%	2.25%	0.65%	1.22%
g_h	0.68%	0.74%	0.76%	0.81%	0.79%	0.84%
$g_{\mathcal{L}}$	0.74%	1.28%				
g_x	0.92%	1.20%	1.33%	1.29%	0.75%	0.97%
g_a	1.24%	1.64%	0.42%	0.62%	0.09%	0.52%
Input share	.42	.42	.76	.68	.90	.65
Obs	50	50	50	50	50	50

Table 12: Variation of Sectoral Growth Accounting Across States

	Farming		Manufacturing		NMNF	
Year	1840-2000	1900-2000	1840-2000	1900-2000	1840-2000	1900-2000
Relative Variance						
σ_y^2	0.0048%	0.0036%	0.0047%	0.0028%	0.0036%	0.0012%
σ_x^2	0.0015%	0.0012%	0.0006%	0.0004%	0.0004	0.0002%
Input share	.32	.34	.13	.16	.11	.18
Relative Variance, Accounting for Covariance						
σ_y^2	0.0048%	0.0036%	0.0047%	0.0028%	0.0036%	0.0012%
$\sigma_x^2+\sigma_{x,a}$	0.0015%	0.0014%	0.0012%	0.0008%	0.0011%	0.0004%
Input share	.31	.38	.26	.28	.31	.37
Comparison of 90/10 Ratios						
g_y^{90}/g_y^{10}	1.87	1.90	2.55	1.85	7.32	1.76
g_x^{90}/g_x^{10}	2.26	2.31	1.68	1.35	1.79	1.51
Input share	1.20	1.22	.66	.73	.24	.86
BDT Average Decomposition						
Input share	.31	.39	.31	.32	.47	.43
Avg. input share	.54	.57	.34	.39	.28	.46
Avg. input share without 90-10 ratio	.31	.36	.24	.27	.29	..33
Obs	50	50	50	50	50	50

Table 13: Sectoral Growth Accounting Across Regions

	Farming			Manufacturing			NMNF		
	North	South	West	North	South	West	North	South	West
1840-2000									
Unweighted									
g_y	2.17%	2.34%	2.39%	1.49%	1.77%	2.33%	1.04%	1.30%	0.62%
g_k	1.72%	2.21%	2.55%	1.86%	2.40%	2.85%	0.86%	1.10%	0.30%
g_h	0.72%	0.73%	0.77%	0.73%	0.77%	0.80%	0.77%	0.81%	0.81%
$g_{\mathcal{L}}$	0.82%	1.03%	1.88%						
g_x	0.95%	1.10%	1.35%	1.11%	1.31%	1.48%	0.80%	0.91%	0.64%
g_a	1.22%	1.25%	1.03%	0.39%	0.46%	0.85%	0.24%	0.39%	-0.02%
Input share	44%	47%	57%	74%	74%	63%	77%	70%	104%
Weighted by labor force									
g_y	2.25%	2.33%	1.73%	1.37%	1.90%	2.24%	1.11%	1.29%	-0.39%
g_k	1.63%	2.15%	1.34%	1.89%	2.60%	3.48%	0.95%	1.10%	-0.62%
g_h	0.70%	0.71%	0.61%	0.73%	0.78%	0.77%	0.77%	0.82%	0.80%
$g_{\mathcal{L}}$	0.74%	1.03%	0.30%						
g_x	0.90%	1.07%	0.70%	1.12%	1.38%	1.66%	0.83%	0.92%	0.33%
g_a	1.35%	1.25%	1.02%	0.26%	0.52%	0.58%	0.28%	0.38%	-0.72%
Input share	40%	46%	41%	81%	73%	74%	75%	71%	-85%
1900-2000									
Unweighted									
g_y	2.54%	3.37%	2.36%	1.74%	2.03%	2.47%	1.39%	1.73%	1.18%
g_k	2.14%	3.38%	2.13%	2.05%	2.36%	2.89%	1.13%	1.45%	0.81%
g_h	0.76%	0.84%	0.81%	0.76%	0.85%	0.86%	0.79%	0.90%	0.87%
$g_{\mathcal{L}}$	1.08%	1.80%	1.32%						
g_x	1.11%	1.55%	1.19%	1.19%	1.35%	1.53%	0.90%	1.08%	0.85%
g_a	1.42%	1.82%	1.18%	0.56%	0.67%	0.94%	0.49%	0.65%	0.34%
Input share	44%	46%	50%	68%	67%	62%	65%	62%	72%
Weighted by labor force									
g_y	2.85%	3.31%	2.10%	1.65%	2.02%	2.21%	1.41%	1.78%	1.19%
g_k	2.24%	3.25%	1.64%	2.04%	2.38%	2.45%	1.17%	1.50%	0.89%
g_h	0.73%	0.82%	0.63%	0.75%	0.88%	0.83%	0.77%	0.91%	0.86%
$g_{\mathcal{L}}$	1.17%	1.77%	0.73%						
g_x	1.13%	1.51%	0.86%	1.17%	1.37%	1.36%	0.91%	1.11%	0.87%
g_a	1.72%	1.80%	1.24%	0.48%	0.64%	0.85%	0.51%	0.68%	0.32%
Input share	40%	46%	41%	71%	68%	62%	64%	62%	73%
Obs	21	16	13	21	16	13	21	16	13

Table 14: Variation of Sectoral Growth Accounting Across States

	Farming			Manufacturing			NMNF		
	North	South	West	North	South	West	North	South	West
1840-2000									
Relative Variance									
σ_y^2	0.0049%	0.0014%	0.0092%	0.0024%	0.0024%	0.0073%	0.0012%	0.0016%	0.0080%
σ_x^2	0.0009%	0.0005%	0.0030%	0.0003%	0.0003%	0.0008%	0.0001%	0.0002%	0.0007%
share	.19	.36	.33	.11	.12	.11	.11	.13	.08
Relative Variance, Accounting for Covariance									
σ_y^2	0.0049%	0.0014%	0.0092%	0.0024%	0.0024%	0.0073%	0.0012%	0.0016%	0.0080%
$\sigma_x^2+\sigma_{xa}$	0.0002%	0.0006%	0.0044%	0.0004%	0.0005%	0.0016%	0.0004%	0.0006%	0.0022%
share	.04	.43	.48	.17	.21	.22	.30	.35	.28
Comparison of 90/10 Ratios									
g_y^{90}/g_y^{10}	1.64	1.38	2.31	2.03	2.42	2.52	2.06	2.13	-
g_x^{90}/g_x^{10}	2.29	1.32	2.92	1.39	1.42	1.63	1.30	1.41	-
share	1.40	.96	1.26	.69	.59	.64	.63	.66	-
BDT Average Decomposition									
share	.09	.43	.49	.18	.23	.28	.45	.49	.46
$\overline{share}$	.43	.55	.64	.29	.29	.32	.37	.41	.27
$\overline{share}_{\text{no 90-10}}$	.10	.41	.43	.15	.19	.20	.29	.33	.27
1900-2000									
Relative Variance									
σ_y^2	0.0021%	0.0006%	0.0027%	0.0015%	0.0009%	0.0043%	0.0004%	0.0012%	0.0007%
σ_x^2	0.0008%	0.0005%	0.0016%	0.0001%	0.0001%	0.0007%	0.0001	0.0002%	0.0001%
share	.36	.87	.59	.10	.06	.16	.18	.18	.16
Relative Variance, Accounting for Covariance									
σ_y^2	0.0021%	0.0006%	0.0027%	0.0015%	0.0000%	0.0043%	0.0004%	0.0012%	0.0007%
$\sigma_x^2+\sigma_{xa}$	0.0004%	0.0003%	0.0014%	0.0003%	0.0005%	0.0010%	0.0002%	0.0005%	0.0001%
share	.16	.53	.53	.24	.04	.24	.37	.40	.14
Comparison of 90/10 Ratios									
g_y^{90}/g_y^{10}	1.56	1.22	1.65	1.63	1.39	2.05	1.43	1.78	1.58
g_x^{90}/g_x^{10}	1.90	1.48	1.84	1.32	1.14	1.53	1.27	1.47	1.31
share	1.22	1.23	1.13	.81	.83	.75	.89	.82	.84
BDT Average Decomposition									
share	.19	.52	.53	.32	.04	.25	.43	.47	.14
$\overline{share}$	.48	.79	.70	.37	.24	.35	.47	.47	.32
$\overline{share}_{\text{no 90-10}}$	.24	.64	.55	.22	.05	.22	.32	.35	.15
Obs	21	16	13	21	16	13	21	16	13

Table 15: Sectoral Growth Accounting By Time Periods

	Farming				Manufacturing			
	1840-80	1880-1920	1920-60	1960-2000	1840-80	1880-1920	1920-60	1960-2000
Unweighted								
g_y	0.51%	0.90%	4.19%	2.74%	-0.83%	2.31%	2.20%	2.24%
g_k	1.59%	1.16%	2.82%	2.43%	0.29%	3.01%	2.12%	2.51%
g_h	0.57%	0.59%	0.80%	0.86%	0.64%	0.67%	0.91%	0.78%
$g_{\mathcal{L}}$	1.26%	0.44%	1.73%	1.51%				
g_x	0.70%	0.68%	1.40%	1.32%	0.52%	1.44%	1.31%	1.35%
g_a	-0.19%	0.22%	2.79%	1.43%	-1.35%	0.86%	0.89%	0.89.%
Input share	1.38	.76	.33	.48	-.64	.63	.60	.60
Weighted by labor force								
g_y	0.89%	0.59%	4.54%	2.65%	-0.08%	1.74%	1.97%	2.24%
g_k	0.62%	1.17%	2.62%	2.43%	1.22%	2.32%	1.96%	2.57%
g_h	0.62%	0.56%	0.79%	0.76%	0.68%	0.63%	0.90%	0.77%
$g_{\mathcal{L}}$	-0.21%	0.06%	1.46%	1.58%				
g_x	0.46%	0.59%	1.30%	1.27%	0.86%	1.19%	1.25%	1.37%
g_a	0.43%	-0.00%	3.25%	1.38%	-0.94%	0.55%	0.72%	0.87%
Input share	.52	1.00	.29	.48	-10.36	.68	.63	.61
Obs	35	45	48	50	35	45	48	50

	NMNF			
	1840-80	1880-1920	1920-60	1960-2000
Unweighted				
g_y	0.74%	-0.00%	1.78%	1.49%
g_k	0.74%	-0.00%	1.88%	0.70%
g_h	0.69%	0.72%	0.97%	0.76%
g_x	0.71%	0.48%	1.27%	0.74%
g_a	0.03%	-0.49%	0.51%	0.75%
Input share	.96	-634.19	.71	.49
Weighted by labor force				
g_y	1.42%	-0.05%	1.81%	1.53%
g_k	1.42%	-0.05%	1.90%	0.76%
g_h	0.74%	0.68%	0.97%	0.76%
g_x	0.96%	0.44%	1.27%	0.76%
g_a	0.45%	-0.49%	0.54%	0.77%
Input share	.68	-8.59	.70	.50
Obs	35	45	48	50

Table 16: Variation of Sectoral Growth Accounting: By Time Periods

	Farming				Manufacturing			
	1840-80	1880-1920	1920-60	1960-2000	1840-80	1880-1920	1920-60	1960-2000
Relative Variance								
σ_y^2	0.0586%	0.0132%	0.0265%	0.0197%	0.0860%	0.0242%	0.0087%	0.0019%
σ_x^2	0.0127%	0.0010%	0.0014%	0.0025%	0.0150%	0.0046%	0.0016%	0.0013%
share	.22	.08	.05	.13	.17	.19	.18	.72
Relative Variance, Accounting for Covariance								
σ_y^2	0.0586%	0.0132%	0.0265%	0.0197%	0.0860%	0.0242%	0.0087%	0.0019%
$\sigma_x^2 + \sigma_{x,a}$	-0.0003%	0.0027%	0.0007%	0.0065%	0.0256%	0.0095%	0.0025%	0.0006%
share	-.00	.21	.03	.33	.30	.39	.29	.33
Comparison of 90/10 Ratios								
g_y^{90}/g_y^{10}	-1.18	-	3.14	9.88	-	4.59	3.27	1.69
g_x^{90}/g_x^{10}	-8.85	-	2.17	3.89	-	3.48	2.24	1.99
share	7.50	-	.69	.39	-	.76	.69	1.18
BDT Average Decomposition								
share	.09	.30	.03	.44	.33	.45	.32	.36
$\overline{share}$	1.95	.79	.19	.32	.27	.65	.37	.65
$\overline{share}_{\text{no 90-10}}$	.10	.19	.04	.31	.27	.62	.26	.47
Obs	35	45	48	50	35	45	48	50

	NMNF			
	1840-80	1880-1920	1920-60	1960-2000
Relative Variance				
σ_y^2	0.1136%	0.0065%	0.0032%	0.0012%
$\sigma_x^2 + \sigma_{x,a}$	0.0112%	0.0010%	0.0005%	0.0003%
share	.10	.15	.14	.25
Relative Variance, Accounting for Covariance				
σ_y^2	0.1136%	0.0065%	0.0032%	0.0012%
$\sigma_x^2 + \sigma_{x,a}$	0.0353%	0.0024%	0.0011%	0.0005%
share	.31	.37	.34	.42
Comparison of 90/10 Ratios				
g_y^{90}/g_y^{10}	-1.10	-	2.47	1.77
g_x^{90}/g_x^{10}	-3.46	-	1.47	1.94
share	3.15	-	.60	1.10
BDT Average Decomposition				
share	.49	.47	.44	.45
$\overline{share}$	1.01	.33	.38	.55
$\overline{share}_{\text{no 90-10}}$	.30	.33	.31	.37
Obs	35	45	48	50

References

Baier, S., Dwyer, G., and Tamura, R. "How Important are Capital and Total Factor Productivity for Growth?" *Economic Inquiry* 44, 2006: 23-49.

Barro, R.J., and Lee, J.-W. "International Comparisons of Educational Attainment," *Journal of Monetary Economics* 32, 1993: 363-394.

Barro, R.J., and Lee, J.-W. "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers* 53, 2001: 541-563.

Berkowitz, D., and Clay, K. "American Civil Law Origins: Implications for State Constitutions and State Courts," *American Law and Economics Review* 7, 2005: 62-84.

Berkowitz, D., and Clay, K. "The Effect of Judicial Independence on Courts: Evidence from the American States," *Journal of Legal Studies* 35, 2006: 399-440.

Berry, W.D., Fording, R.C., and Hanson, R.L. "An Anual Cost of Living Index for the American States, 1960-95," *Journal of Politics* 60, 2000: 550-567.

Bils, M., and Klenow, P.J. "Does Schooling Cause Growth?" *American Economic Review* 90, 2000: 1160-1183.

Bleakley, H. "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics* 122, 2007: 73-117.

Bleakley, H. "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure," *American Economic Journal: Applied Economics* 2, 2010: 1-45.

Bodenhorn, H. *State Banking in Early America: A New Economic History*. Oxford University Press, 2002.

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

Caselli, F. "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, eds. by P. Aghion, and S.N. Durlauf, vol. 1A, chapter 9, 679-741, Elsevier, 2005.

Caselli, F., and Coleman, W.J. "The U.S. Structural Transformation and Regional Convergence," *Journal of Political Economy* 109, 2001: 584-616.

Easterlin, R. "Interregional Differences in Per Capita Output, Population and Total Output, 1840-1950," in *Trends in the American Economy in the Nineteenth Century*, ed. by W.N. Parker, Princeton University Press, 1960a.

Easterlin, R. "Regional Growth of Output: Long Term Tendencies, 1880-1950," in *Population and Redistribution and Economic Growth, United States, 1870-1950*, ed. by S. Kuznets, A.R. Miller, and R. Easterlin, vol. 1 of *Analyses of Economic Change*. American Philosophical Society 1960b.

Easterly, W., and Levine, R. "It's Not Factor Accumulation: Stylized Facts and Growth Models," World Bank, 2002.

Gallman, R.E. "The United States Capital Stock in the Nineteenth Century," in *Long-Term Factors in American Economic Growth*, ed. S.L. Engerman, and R.E. Gallman, University of Chicago Press, chap 4, 165-213, 1986.

Garofalo, G.A., and Yamarik, S. "Regional Convergence: Evidence from a New State-By-State Capital Stock Series," *Review of Economics and Statistics* 84, 2002: 316-323.

Gordon, R. *Macroeconomics*. Addison-Wesley, New York, 1999.

Hall, R.E., and Jones, C.I. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114, 1999: 83-116.

Heston, A., Summers, R., and Aten, B. *Penn Word Table Version 6.3*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Jorgenson, D.W. "Accounting for Growth in the Information Age," in *Handbook of Economic Growth*, eds. P. Aghion, and S.N. Durlauf, vol. 1A, chapter 10, 743-815, Elsevier, 2005.

Jorgenson, D.W., and Gollop, F.M. "Productivity Growth in U.S. Agriculture: A Postwar Perspective," *American Journal of Agricultural Economics* 74, 1992: 745-750.

Jorgenson, D.W., Gollop, F.M., and Fraumeni, B.M. *Productivity and U.S. Economic Growth.*, Harvard University Press: Cambridge, 1987.

Klenow, P.J., and Rodriguez-Clare, A. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in *NBER Macroeconomics Annual 1997*, eds., by B.S. Bernanke, and J.J. Rotemberg, MIT Press, 1997: 83-103.

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

Mitchener, K.J., and McLean, I.W. "U.S. Regional Growth and Convergence 1880-1980," *Journal of Economic History* 59, 1999: 1016-1042.

Mitchener, K.J., and McLean, I.W. "The Productivity of U.S. States Since 1880," *Journal of Economic Growth* 8, 2003: 73-114.

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

Romer, P.M. "Increasing Returns and Long-Run Growth," *Journal of Political Economy* 94, 1986: 1002-1037.

Romer, P.M. "Endogenous Technological Change," *Journal of Political Economy* 98, 1990: S71-S102.

Ruggles, S., Alexander, J.T., Genadek, K., Goeken, R., Schroeder, M.B., and Sobek, M. "Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]," Minneapolis: University of Minnesota, 2010.

Schoellman, T., and Tamura, R. "Development Accounting and Convergence for US States," 2011, Arizona State University working paper.

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

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

Tamura, R., Dwyer, G.P., Devereux, J., and Baier, S.L. "Economic Growth in the Long-Run," 2011 Clemson University working paper.

Tostlebe, A.S. *The Growth of Physical Capital in Agriculture, 1870-1950*. UMI, 1954.

Turner, C., Tamura, R., Mulholland, S.E., and Baier, S.L. "Education and Income of the States of the United States: 1840-2000," *Journal of Economic Growth* 12, 2007: 101-158.

Turner, C., Tamura, R., Schoellman, T., and Mulholland, S.E. "Estimating Physical Capital and Land for States and Sectors of the United States, 1850-2000," Arizona State University working paper, 2010.

Weil, D. N. *Economic Growth* second edition, Pearson Addison Wesley, 2009.

Weiss, T. “Estimates of White and Nonwhite Gainful Workers in the United States by Age Group and Sex, 1800 to 1900,” *Historical Methods* 32, 1999L 21-36.

Williamson, J.G. and Linder, P.H. *American Inequality* Academic Press, 1980.

World Development Indicators 2006.

Young, A. “The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience,” *Quarterly Journal of Economics* 110, 1995: 641-680.

Young, A. “Gold into Base MetalsL Productivity Growth in the Peoples Republic of China during the Reform Period,” *Journal of Political Economy*— 111, 2003: 1220-1261.

6 Data Availability

Our series for inputs and output do not span the entire 1840-2000 period for each state. The year in which a complete record of inputs and outputs for each state becomes available is as follows:

1840: Alabama, Arkansas, Connecticut, Delaware, Florida, Georgia, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Mississippi, Missouri, New Hampshire, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Vermont, Virginia

1850: California, New Mexico, Oregon, Texas, Utah, Wisconsin

1860: Kansas, Minnesota, Nebraska, Washington, Utah

1870: Colorado, Montana, Nevada, West Virginia

1880: Arizona, Idaho

1900: North Dakota, South Dakota, Wyoming

1920: Oklahoma

1930: District of Columbia

1950: Alaska, Hawaii.

Specific series may be available earlier.

7 Regions and Subperiods

7.1 Regions

Reference in the paper is often made to the North, South, and West. We divide the states so that the West comprises roughly the arid states that benefited most from irrigation and the South comprises roughly those states below the Mason-Dixon line. Each categorization includes at least two Census regions, listed in italics. The exact breakdown is given in Table 17.

Table 17: Division of States into Regions

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

* D.C. is excluded from the empirical analyses. This is due to the special nature of its production: it is a city, and overwhelmingly its output after World War II is in government services.