Regional Convergence and Technological Progress in American Manufacturing before World War II

Matthew Hurt

March 23, 2022

1 Introduction

The first half of the twentieth century is often seen as an era of increasing regional labor market integration (Wright, 1999; Rosenbloom, 2002) measured by a convergence in regional real wages. While the migration of workers drove convergence by changing regional labor supply, changes in the regional labor demand from firms also drove the convergence in real wages. As initially poorer regions of the country industrialized the emergence on national trusts and new national regulations of industrial activity may have caused relative greater gains in productivity and lead to relatively higher wage growth.

In other words, convergence in manufacturing income (which mirrored the overall economy) may have been driven by the same kind of industrial activity becoming more similar across states, or the manufacturing activity within each state may have been become more similar across all industry groups. Identifying which force played the more important role in promoting convergence is important to understand the nature of America's 20th century economic convergence. Economists interested in understanding the comparative economic development of contemporary countries will also want to better understand these forces because there are serious parallels between the regional dispersion of income in the US historically and across countries today.

This paper uses new data from the United States Census of Manufactures between 1900-1947 to demonstrate that convergence was present in and faster for the manufacturing sector of the economy, and measures whether convergence across industry groups or across states was stronger within manufacturing. The main advantage of these new data are that they allow for consistent comparisons across time, states, and industry groups. Using the full model with year, state, and industry group fixed effects I find a convergence coefficient greater than 10%, which is significantly higher than previous estimates in the 2-6% range found in the literature for the entire economy. Aggregating the economy to the state level I find the rate of convergence remains virtually unchanged. Aggregating the economy to the industry group level I find that the rate of convergence falls by two percentage points. These results demonstrate that both supply and demand forces drove convergence, but that convergence was stronger across states than across industry groups.

The list of papers that have studied convergence in the American economy is long (Barro and Sala-i-Martin, 1992; Bernard and Jones, 1996; Evans and Karras, 1996b,a; Evans, 1997; Higgins et al., 2006; Young et al., 2008) and have typically focused on state or county level GDP estimates and included some sectoral analysis. To my knowledge Barro and Sala-i-Martin (1991) is the only paper to study the manufacturing sector of the historical American economy though only as one of many sectors. Manufacturing has been identified as one of the most important sectors of the economy with respect to convergence Rodrik (2013). Even skeptics have found that growth in manufacturing productivity can lead to spillovers into other sectors of the economy which foster convergence. Herrendorf et al. (2022)

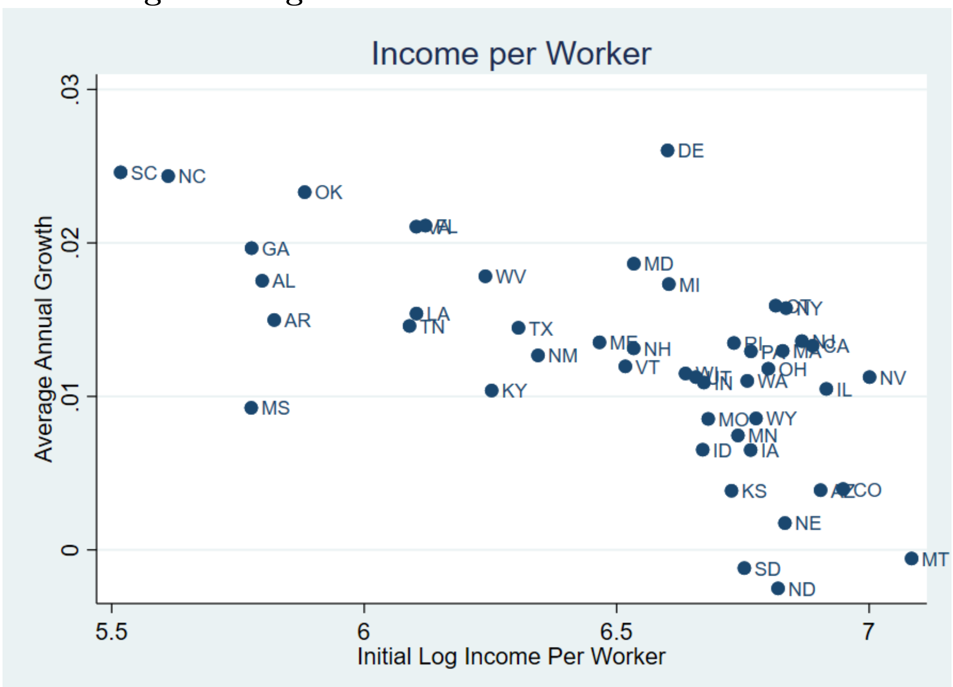
The United States during this period also serves as a valuable out of sample review for a phenomenon identified by Rodrik (2013). That paper identified that even when convergence failed to take place in cross-country analysis convergence in manufacturing was taking place. The reason for this discrepancy was that poorer nations also tended to have smaller manufacturing sectors and the convergence in manufacturing was not sufficient to offset the lack of convergence in non-manufacturing.

Instead the United States exhibited both convergence in income overall and a faster rate of convergence in manufacturing. However new evidence shows that cross country convergence is now taking place even while it failed to during the 20th century (Kremer et al., 2021). This makes the United States during the 20th century which experienced both convergence in income and convergence in manufacturing income a useful period to study.

Figure 1 presents a long view of convergence in income per worker by scattering the average annual growth in state income per worker between 1900 and 1940 against the log of each state's income per worker in 1900. The data show that poorer states, based on the 1900 distribution of state income per worker, grew faster than wealthier states. This negative relationship is visual evidence of convergence in income per worker.

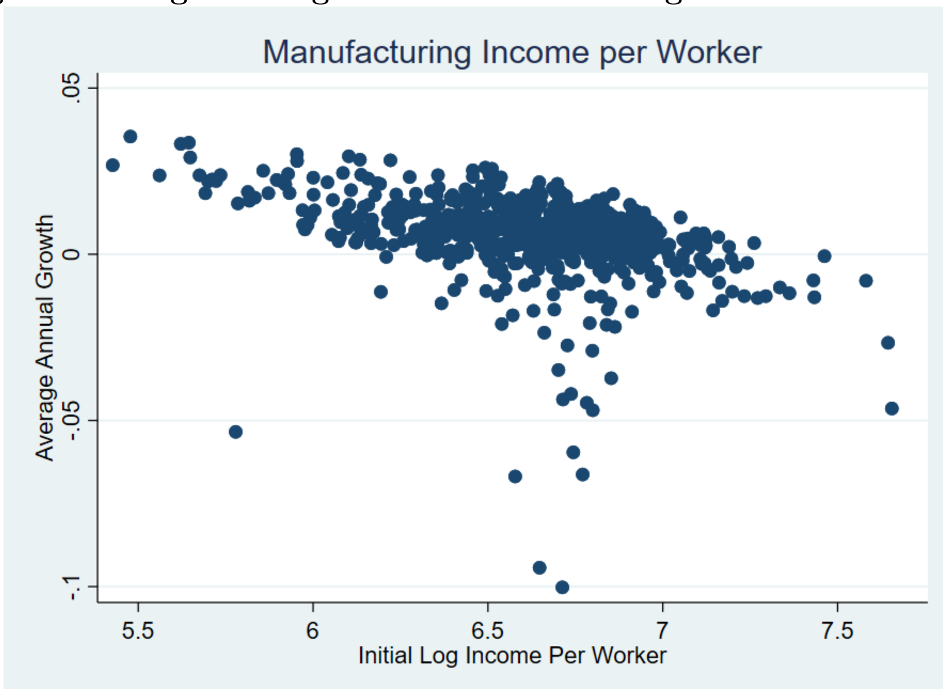
Figure 1 can also help us think about what faster or slower convergence looks like. A steeper slope indicates a faster rate of convergence. When initial incomes per worker are held constant, but

Figure 1: Long Convergence in Income Per Worker at the State Level



Note: This figure scatters the average annual growth rate in income per worker between 1900-1940 against 1900 log income per worker at the state level. The data used for this figure come from a variety of sources including Klein (2013); Steven Ruggles and Sobek. (2021); Easterlin (1960) along with data from the Bureau of Economic Affairs.

Figure 2: Long Convergence in Manufacturing Income Per Worker



Note: This figure scatters the average annual growth rate in income per worker between 1900-1940 against 1900 log income per worker using data from the US Census of Manufactures.

growth rates for the poorest states rise while growth rates for the wealthiest states decline this leads to a faster rate of convergence. Convergence can play a role in explaining the pattern of differential growth in state income over time, but it is important to keep the concept separate from economic growth. While Montana and South Carolina would likely both be happy to have higher growth and not begrudge the other start for experiencing a high rate of growth, South Carolina was greatly aided by the process of convergence that existed in the United States between 1900 and 1940 in a way Montana did not. This is important to keep in mind when exploring patterns of converge later in the paper.

Figure 2 scatters the average growth in manufacturing income per worker for each state-industry group pair between 1900 and 1940 against the initial manufacturing income per worker for each state-industry group in 1900. As was the case in Figure 1 the clear negative relationship indicates that convergence was also present in manufacturing.

The similar bounds on the x-axis demonstrate that in the United States manufacturing income per worker was not necessarily higher than overall state level income per worker. Part of this can be explained by large manufacturing shares of the overall labor force, but it is also the case that in many states non-manufacturing activities often had higher levels of income per worker.

Figure 2 also demonstrates that growth in manufacturing could be faster than in the economy overall, but certain state manufacturing industry pairs experienced significant decline in manufacturing income per worker. These outliers underscore the importance of using an empirical strategy that can control for decadal macroeconomic shocks as well as state and industry group level features that inform the process of convergence.

2 Empirical Strategy

Convergence in income across economies is a prediction that follows from neoclassical growth models, the simplest of which is the Solow model. When there is no growth in population or technology the model predicts a steady state level of capital per worker dependent on economy level parameters such as the level of technology, or savings rate. The concave nature of the per worker production function implies that economies that are further below their steady state level of capital per worker will move more quickly toward their steady state. In the most extreme case where every economy shares the same parameters an economy that starts with a lower level of capital per worker will experience faster growth in capital per worker and income per worker.

In the literature the most frequent test of this relationship is described by equation (1) where an average growth rate in income per worker is regressed on a constant term and the initial value of income per worker.

$$\hat{y}_i = \beta_0 - \beta ln(y_i^t) + \epsilon_i \tag{1}$$

In equation (1) $\hat{y}_i = \frac{1}{T-t+1}(ln(y_i^T) - ln(y_i^t))$ where y_i is real income per worker for an economy i with initial time period t and ending time period T. β is a pseudo elasticity and the measure of convergence in income per worker for an economy. [1] Fixing t, if the magnitude of β increases when equation (1) is estimated using T instead of T' for the end year then the rate of convergence was faster between t and T. Faster convergence means that economies that begin period t poorer will experience relatively faster growth.

This approach is convenient in that it only requires two time periods for each iteration and is analogous to Figure 1. Nominal values can be used in this equation although one should be careful that in those situations the observed convergence is not just in price levels. While this empirical strategy has yielded important insight into convergence across states and countries, it has been improved upon in different ways. One natural approach is to add control variables, such as measures of technology, institutional quality, and others that can help explain the sources of convergence. This approach is more common in cross-country analysis rather than analysis that focuses on the United States since the states within a country are assumed to have more of these features in common than countries will.

The convergence described by equation (1) is a form of *long* convergence because the outcome variable is an average annual growth rate for one time period. These long averages are useful, but obscure short run economic phenomena that can be relevant to understanding patterns of convergence. In contrast *stacked* convergence involves simultaneously estimating a convergence coefficient using information on many shorter time periods.

$$\hat{y_{it}} = \beta_0 - \beta ln(y_{it}) + \gamma_t + \epsilon_{it} \tag{2}$$

The difference between equation (1) and (2) is that now β is estimated simultaneously for each time period and the average growth rate between each period. For example, suppose the data cover three different years. In that case a researcher can use equation (1) to run three separate regressions to estimate β between the first two years, the second and third, and then the long measure from the first to the third. But, a researcher can also include the initial values from year one paired with the

[1] In the literature it is common to study income per capita instead of income per worker. Both are interesting and may converge in different ways, but since the manufacturing data are at the per worker level this paper also uses income per worker for the entire economy.

average annual growth rate from year one to year two and the initial values from year two paired with the average annual growth rate from year two to year three to simultaneously estimate β.

One of the primary advantages of this *stacked* approach is that it allows for the inclusion of year fixed effects. While there is a debate about the value of including year fixed effects in growth regressions (Acemoglu et al., 2005), the inclusion of the fixed effects allows for a measure of convergence taking place in spite of time period specific economic shocks. For the United States between 1900 and 1947 which experienced many economic shocks including two world wars and the Great Depression, a model like equation (2) is preferable.

When conditioning variables are included the interpretation of β changes. Instead of a measure of the *unconditional* convergence in income per worker β now measures how much faster the poorer version of an economy with the conditioned characteristics will grow relative to a version of the same economy with a larger initial income per worker. For example, when country fixed effects are included β can be used to estimate how quickly the country will reach its income per worker frontier. [2]

$$\hat{y_{ijt}} = \beta_0 - \beta ln(y_{ijt}) + \gamma_i + \gamma_j + \gamma_t + \epsilon_{ijt} \tag{3}$$

In equation (3) there are now separate fixed effects included for each region i and for an additional characteristic j. In this paper i are US states and j are industry groups. The γ_i capture time invariant features of states while the γ_j capture time invariant features of each industry group that are relevant to the process of convergence. It is well known in the literature that β increases when conditioning variables are included (Rodrik, 2013), which is not surprising since in the context of a Solow model it would be trivial to explain the pattern of cross country economic development if the model were calibrated to each individual country. Each of the above estimation strategies have their merits and are used in this paper. Descriptions and results tables will clarify which approach is used to generate which results.

[2] Specifically, for some initial share of a country's frontier α, β is used to find the new share of a country's frontier after T years have passed using the following formula: $\alpha^{(1-\beta)^T}$

3 Data

The primary data used in this paper come from the United States Census of Manufactures for the years 1900, 1910, 1920, 1930, 1940, and 1947. These reports include the number of establishments, wage earners, salaried workers, along with wage bills for wage earners and salaried workers, as well as value, and value added for all sizable industrial activity within each state measured in nominal values. The state level values were hand recorded by the author. The industrial activity was then assigned one of twenty SIC-2 industrial code to harmonize industrial activity across the six periods.

The Census of Manufactures includes industrial activity for the 48 continental United States and D.C. in each year. However, because the national data exclude D.C. and because the data for the state of Oregon are missing in 1910 I exclude both D.C. and Oregon from all analysis. This leaves 47 states and twenty industry groups so that the total number of state-industry group pairs in any given year is 940 though there are some null values since not all states engaged in all types of industrial activity.

Manufacturing income per worker is computed by summing the reported wage bill for salaried earners with the reported wage bill for wage earners and then dividing by the sum of both worker types. Growth rates in income per worker are computed for 1900 to 1910, 1910 to 1920, 1920 to 1930, 1930 to 1940, 1940 to 1947 and from 1900 to 1940 for the long analysis. Thus the final number of total possible observations used in the baseline analysis is 4700. However, null values are present for reasons previously described.

Table 1 presents summary statistics of key trends in aggregate manufacturing between 1900 and 1947. Column 4 highlights the emergence of larger establishments - a trend that was only paused by the Great Depression. This trend happened both because the manufacturing labor force was generally increasing, but also because the total number of establishments declined. The 1947 Census of Manufactures notes that the period between 1940-1947 was an unprecedented period of expansion in manufacturing activity and prosperity which explains the increase in establishments. The composition of the type of work done in manufacturing was also changing in this period. In the historical setting salaried work is equated with higher skill work so column 5 highlights the

Table 1: Annual Descriptive Statistics

(1) Year	(2) Establishments	(3) Manufacturing Labor Force	(4) Employees Per Establishment	(5) Salaried Share of Employment	(6) Average Salary	(7) Average Wage
1900	515,231	5,753,072	11.17	0.07	1009	439
1910	266,235	7,377,961	27.71	0.11	1188	517
1920	290,114	10,545,129	36.35	0.14	2006	1158
1930	210,964	10,190,256	48.30	0.13	2646	1316
1940	187,740	9,163,796	48.81	0.10	1797	1074
1947	240,881	14,255,703	59.18	0.17	3926	2525

Note: This table presents summary statistics from the US Census of Manufactures. These values include the sum of establishments, sum of wage earners and salaried manufacturing workers, the number of these employees per establishment, the share of that work force engaged in salaried work, and average salaries and wages in nominal terms.

increasing skill intensification of manufacturing. At the same time columns 6 and 7 present some evidence of the list the nominal average salary and wage which reveal convergence in earnings between the two worker types.

When measured in nominal values one possible explanation for observed convergence in income per worker is a convergence in price levels, rather than real income per worker. Because all of the values in the Census of Manufactures are recorded in nominal terms and do not correct for regional price differences I apply corrections before using these data for regression analysis. To control for changes in price level over time I use 1940 equivalent income. [3]

To correct for regional price differences I used data from Table 1 of Mitchener and McLean (1999). In that table division price levels relative to the national average can be calculated for 1900, 1920, and 1940 for each of the nine census divisions. I then linearly interpolate price levels for each division in 1910 and 1930. I use 1940 values for 1947 rather than do a partial interpolation between 1940 and 1960. Once these regional price levels are computed I make adjustments to inflate or deflate regional prices to the national average. For example, if prices in New England were found to be 90% of the national price level in 1900 the income in the Census of Manufactures for New England in 1900 will be increased 11% before being used in regression analysis. Table A1 in the

[3]Estimates of national price levels come from the Minneapolis Federal Reserve. The price levels used are 25, 28, 56, 50, 42, 63 for the years 1900, 1910, 1920, 1930, 1940, and 1947 respectively.

Appendix lists the values that were multiplied against each nominal value. Division level prices are assumed to be the same for all states within each division. Together these corrections allow for the first comparison of real purchasing power parity adjusted manufacturing income per worker across space and time for the US during this period.

4 Results

This section presents estimates of convergence coefficients using different samples and estimation strategies. All of the tables highlight rates of convergence much larger than the usual estimates in the 2-6% range found for the entire economy and greater than the 4% value measured for the manufacturing economy. The differences may seem small, but represent substantial differences over the period studied. An unconditional convergence coefficient of 2% implies that a state that is 75% as wealthy in log points as the frontier state in 1900 will be 89.5% as wealthy in log points as the frontier state in 1947. An unconditional convergence coefficient of 4% implies that same state would be 95.9% of the frontier while an unconditional convergence coefficient of 10% would be fast enough for a state to nearly catch the frontier.

Table 2 presents the first set of results with the β reported in columns (1) - (3) estimated using equation (3) and the β reported in columns (4) - (8) coming from a version of equation (3) with no time dimension or time fixed effects. All observations are weighted by manufacturing labor force size. Only 1940 equivalent region adjusted prices were used in estimation and all standard errors are clustered at the state level.

Column (1) shows an unconditional rate of convergence of 4.5% across all state-industry group pairs between 1900 and 1947. While state industry groups are not states the rate of convergence across all state industry groups is comparable to the rate of convergence across states. This 4.5% convergence rate is substantially faster than the common 2% rate of convergence found in the overall economy. The differences are not trivial. An unconditional rate of convergence of 2% implies that a state industry group with half the log income per worker of the frontier income in 1900 would

Table 2: Convergence in Manufacturing Income Per Worker

Average Growth	(1) Full Sample	(2) Full Sample	(3) Full Sample	(4) 1900-1910	(5) 1910-1920	(6) 1920-1930	(7) 1930-1940	(8) 1940-1947
Income Per Worker	-0.045	-0.065	-0.101	-0.042	-0.028	-0.027	-0.077	-0.137
	(0.007)	(0.009)	(0.012)	(0.007)	(0.005)	(0.008)	(0.017)	(0.005)
Year FE	Yes	Yes	Yes	No	No	No	No	No
State FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Observations	3493	3493	3493	720	711	700	683	679
R-squared	0.663	0.704	0.769	0.549	0.715	0.608	0.292	0.936

Note: This table presents results from eqn. (3) using data from the US Census of Manufactures between 1900 and 1947. The outcome variable is the decadal growth rate of income per worker while the independent variable is the log of income per worker in the beginning year for each year pair. All values are computed using 1940 equivalent census division adjusted prices and are weighted by labor force. Standard errors are clustered at the state level.

have 76.5% of the frontier in 1947, while a rate of 4.5% implies that the same state industry group would be 92% of the frontier in 1947.

The remaining columns list conditional convergence rates. Columns (2) and (3) show the rates of conditional convergence across all industry group state pairs between 1900 and 1947. The 6.5% reported in column (2) implies that a state that starts at 50% of its frontier in log points, determined by the industrial composition of the state, would reach 97% of its frontier after 47 years. The 10% convergence rate from column (3), which now includes state fixed effects, has the interpretation that a state industry group that begins at 50% of its potential would reach that same 97% threshold in only thirty years and would effectively reach its frontier after 47 years.

The remaining columns (4) - (8) highlight how convergence changed over time. By estimating each coefficient separately we are computing five *long* convergence coefficients rather than one *stacked* one. Visually this is equivalent to producing Figure 2 for each of the five time periods. The results show a striking variability over time with convergence in manufacturing operating at the same speed around the generally accepted measure for the overall economy between 1910-1930. Convergence accelerates dramatically after 1930.

While the exact values calculated here are not comparable to the existing decadal convergene estimates from Barro and Sala-i-Martin (1992) the time trends are. They also measure conditional convergence coefficients by conditioning on a sectoral concentration variable similar to an HHI used to measure market concentration. Their results cover a broader time period, from 1880-1988, but are for the entire economy rather than for manufacturing. As in this paper they find that the period between 1940-1950 exhibited the fastest rate of convergence and that convergence was typically stronger before 1900 than it was after the turn of the century. The major difference is that they find that the period between 1930-1940 exhibited the slowest rate of convergence.

The timing matters. The results in this paper can be used to argue that the Great Depression and/or New Deal policies were instrumental in promoting a faster rate of convergence. For example, if New Deal policies forced more efficient practices on the poorest state industry pairs in 1930 that could explain why these poorer states grew relatively quickly. The rapid convergence between 1940 and 1947 may be explained by patterns of military spending. If the government targeted the poorest state industry pairs in 1940 of the country for rapid industrialization for war manufacturing that would have likely caused an acceleration in convergence.

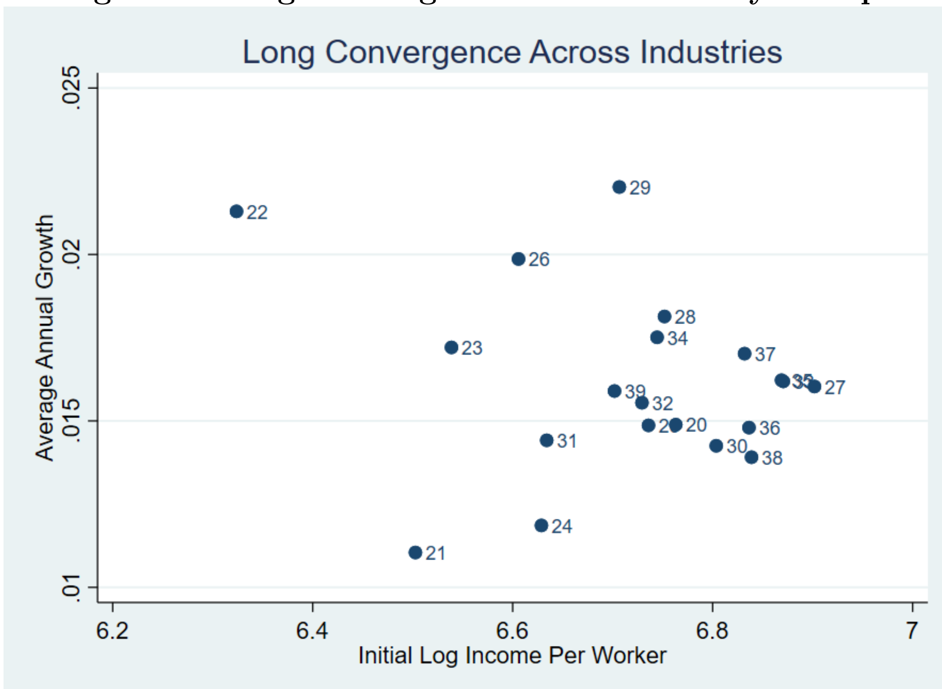
These patterns merit further study, but it can be difficult to interpret convergence coefficients estimated by studying state industry pairs. The remainder of this section will instead focus on estimating the same collection of unconditional and conditional convergence coefficients for states and industries separately. This new analysis will allow us to determine to what extent the convergence found in Table 1 was driven by a convergence in industrial activity of states or industry groups.

4.1 Manufacturing Convergence Across Industry Group

The analysis in this section begins with the same data used in regression analysis for Table 2. These data are then collapsed into measures of industry group level manufacturing activity by summing across the forty seven states for each year. Labor force weights are then recomputed at the industry group level. The same 1940 equivalent region adjusted prices are used in estimation. All standard errors are clustered at the industry group level.

Figure 3 presents visual evidence of long convergence across industry groups. These initial values are the 1900 log of total manufacturing income per worker for each industry group while the growth

Figure 3: Long Convergence Across Industry Groups



Note: This figure scatters the average annual growth rate in manufacturing income per worker between 1900-1947 against 1900 log manufacturing income per worker using data from the US Census of Manufactures for each of the SIC 2 industry groups.

rate is computed using 1947 as the end year. Appendix Table A2 lists the industry groups. The downward sloping nature of the data indicate the presence of convergence. The industry group with the lowest earnings per worker in 1900, textiles, grew at the second fastest rate, beyond that the data are complicated. Industries with initial income around 6.6 log points actually appear to exhibit divergence across time, but industry groups with more than 6.7 log points of initial earnings exhibit a negative relationship.

Table 3 presents regression results for a regression strategy similar to equation (3), but omitting the state dimension. Column (1) lists the unconditional convergence coefficient estimated jointly across all time periods. Convergence is present, but relatively weak. The estimate is statistically significant and the process of convergence explains 91.4% of the observed variation in decadal growth rates in manufacturing income per worker, but the 1.6% is below most estimates of convergence across the entire economy. An industry group at 50% of the log points of the frontier industry group in 1900 would only be 72% of the frontier by 1947.

Table 3: Convergence in Manufacturing Across Industry Groups

Average Growth	(1) Industries	(2) Industries	(3) 1900-1910	(4) 1910-1920	(5) 1920-1930	(6) 1930-1940	(7) 1940-1947
Income Per Worker	-0.016	-0.084	-0.005	-0.021	0.010	0.003	-0.048
	(0.004)	(0.015)	(0.004)	(0.011)	(0.010)	(0.011)	(0.008)
Year FE	Yes	Yes	No	No	No	No	No
Industry Group FE	No	Yes	No	No	No	No	No
Observations	100	100	20	20	20	20	20
R-squared	0.914	0.948	0.032	0.223	0.039	0.005	0.721

Note: This table presents the results from eqn. (3) without the state dimension by collapsing all data for each year to the industry group level. The data come from the US Census of Manufactures between 1900 and 1947. The outcome variable is the decadal growth rate of income per worker while the independent variable is the log of income per worker in the beginning year for each year pair. All values are computed using 1940 equivalent census division adjusted prices and are weighted by labor force. Standard errors are clustered at the industry group level.

The conditional rate of convergence is substantially faster and in line with the rates estimated in Table 2 columns (2) and (3). An industry group at 50% of its frontier in log points in 1900 would be at 99% of its frontier by 1947 under this rate of convergence. Decade by decade unconditional convergence rarely seems to explain or be present except between 1940 and 1947. The low observation count makes it difficult to have precisely estimated coefficients in columns (3) - (7), but even the point estimates tend to be slow compared to the convergence across state industry group pairs. The large gap between the estimates in column (1) and (2) indicates that while industry groups that initially offered relatively low wages in 1900 only slowly caught up to the ones that offered initially high wages, industry groups rapidly caught up to their own within industry group frontier wages.

If we assume that labor markets were reasonably competitive such that the differences in wages offered by industry groups were the product of differences in productivity across industry groups then, per column (1), the least efficient industry groups in each year did exhibit faster productivity growth given the convergence rate of 1.6%. However, the rate that industry groups arrived at their own frontier productivity level was dramatically faster for those who were initially further from their productivity frontier. This may indicate that there were few opportunities for industry groups to take advantage of methods or technologies being used in other ones. If that is true, it seems less reasonable that the entry or expansion of high productivity industries into a state would have been responsible for spurring growth.

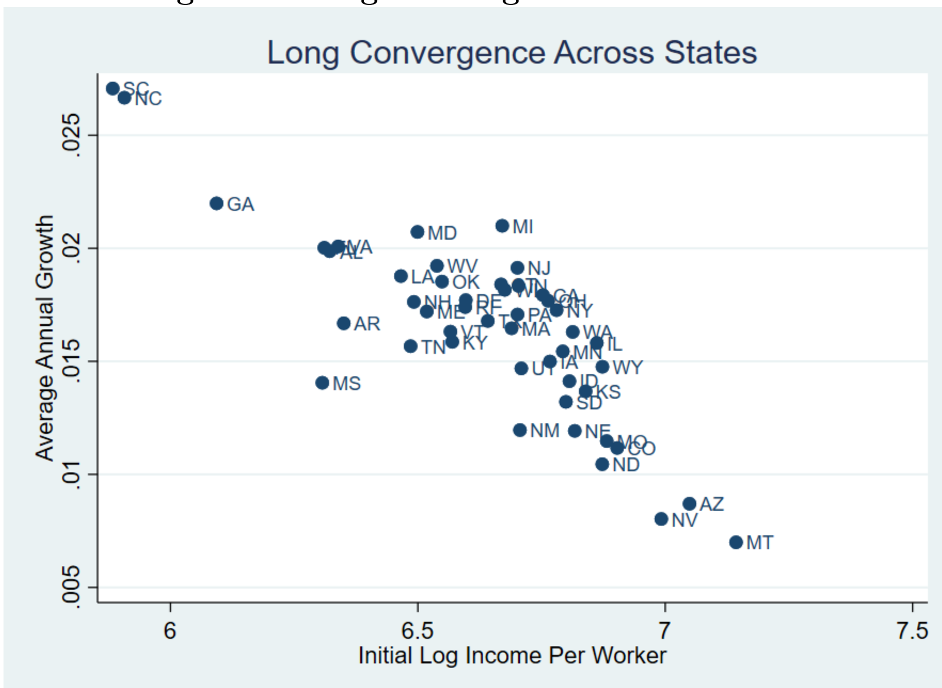
4.2 Manufacturing Convergence Across States

This section repeats the analysis performed above, but across states rather than industry groups. Figure 4 shows a more convincing case for the existing of long convergence between 1900 and 1947 than what was found in Figure 3. Because this figure includes 1947 and the period of 1940-1947 exhibited strong convergence it is difficult to compare the growth rates of Figure 4 to Figure 1, but we can at least see that state level manufacturing income in 1900 tended to be more compressed than state level income was. We can also see that manufacturing income positively correlated with overall state level income.

Table 4 lists the conditional and unconditional convergence coefficients for the continental United States minus Oregon and D.C. Unlike for the industry groups the estimates are statistically significant for all years except 1930-1940. Both the unconditional and conditional convergence coefficients from columns (1) and (2) are faster than the ones found in Table 3 with the conditional coefficient being as fast as the one found for all state industry group pairs in Table 2. The 2.4% unconditional convergence coefficient means that a state with half the log income per worker points of the frontier state in 1900 will be have 80% of those points by 1947. At 10.6% for the conditional convergence coefficient the same state will have effectively reached its frontier in 47 years.

Across states the time trend is similar to the one found in Table 2 for state industry group pairs. The fastest rate of convergence is found between 1940-1947 and the slowest is found between 1920-1930. The differences are that states experienced a precisely estimated and non-trivial divergence in manufacturing income per worker between 1920 and 1930. At a 3% divergence rate a state starting at 50% of the frontier would only be 40% of the frontier after one decade. The lack of convergence between 1930-1940 is also different though it mirrors the findings of Barro and Sala-i-Martin (1992)

Figure 4: Long Convergence Across States



Note: This figure scatters the average annual growth rate in manufacturing income per worker between 1900-1947 against 1900 log manufacturing income per worker using data from the US Census of Manufactures for each of the 47 states included in this analysis.

Table 4: Convergence in Manufacturing Across States

Average Growth	(1) Year	(2) Both	(3) 1900-1910	(4) 1910-1920	(5) 1920-1930	(6) 1930-1940	(7) 1940-1947
Income Per Worker	-0.024	-0.106	-0.015	-0.033	0.03	-0.003	-0.059
	(0.003)	(0.015)	(0.003)	(0.005)	(0.011)	(0.010)	(0.011)
Year FE	Yes	Yes	No	No	No	No	No
State FE	No	Yes	No	No	No	No	No
Observations	235	235	47	47	47	47	47
R-squared	0.902	0.946	0.272	0.389	0.235	0.004	0.709

Note: This table presents the results from eqn. (3) without the industry group dimension by collapsing all data for each yaer to the state level. The data come from the US Census of Manufactures between 1900 and 1947. The outcome variable is the decadal growth rate of income per worker while the independent variable is the log of income per worker in the beginning year for each year pair. All values are computed using 1940 equivalent census division adjusted prices and are weighted by labor force. Standard errors are clustered at the state level.

who found a slow rate of unconditional convergence between 1930-1940 and divergence between 1920 and 1930.

The same substantial difference between unconditional and conditional coefficients in columns (1) and (2) is present as in Table 3. This difference means that states were faster at reaching their own frontiers than reaching the frontier state. At this point it is tempting to pivot to regional analysis and see if the same trends hold (within a region are the gaps narrower?). The usual means of doing sub-sample analysis involve running the primary regression removing regions one at a time. Performing this kind of analysis at the census division level reveals no statistically significant differences in convergence coefficients.

The coefficients in Table 4 are always larger in magnitude than the ones in Table 3, except for column (6), meaning that the forces of convergence were faster in states than across industry groups. Increasing similarities within and across industry groups drove some of the observed convergence across state industry group pairs, but the data support the notion that convergence was dominated by cross-state forces. In other words, states were beginning to have more similar income per worker due to changes in other parts of the economy that would have affected the manufacturing wage.

Lastly, this is the first paper to perform these kind of analysis for the United States using 1940 region adjusted prices. In previous studies some of the convergence was undeniably the result of a convergence in regional price levels. The results in this paper, including the important result that convergence across states outpaced convergence in industry groups holds when using nominal data. Analysis involving the removal of the top and bottom 5/10% outliers also does not change the nature of the results though it does globally reduce the magnitude of the coefficients.

References

Acemoglu, D., S. Johnson, and J. Robinson (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 385–472.

Barro, R. and X. Sala-i-Martin (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, 107–182.

Barro, R. and X. Sala-i-Martin (1992). Convergence. *The Journal of Political Economy 100*(2), 223–251.

Bernard, A. and C. Jones (1996). Productivity and convergence across u.s. states and industries. *Empirical Economics 21*(1), 113–135.

Easterlin, R. (1960). Regional growth of income: Long term tendencies. In *In Population Redistribution and Economic Growth, United States, 1870-1950, vol. 2, Analysis of Economic Change.* edited by Simon Kuznets, Ann Ratner Miller, and Richard A. Easterlin: American Philosophical Society.

Evans, P. (1997). How fast do economies converge? *The Review of Economics and Statistics 79*(2), 219–225.

Evans, P. and G. Karras (1996a). Convergence revisited. *Journal of Monetary Economics 37*, 249–265.

Evans, P. and G. Karras (1996b). Do economies converge? evidence from a panel of u.s. states. *The Review of Economics and Statistics 78*(3), 384–388.

Herrendorf, B., R. Rogerson, and Valentinyi (2022, March). New evidence on sectoral labor productivity: Implications for industrialization and development. Working Paper 29834, National Bureau of Economic Research.

Higgins, M., D. Levy, and A. Young (2006). Growth and convergence across the u.s.: Evidence from county-level data. *The Review of Economics and Statistics 88*(4), 671–681.

Klein, A. (2013). New state-level estimates of personal income in the united states, 1880–1910. In *Research in Economic History*. Clark, Gregory, and Sundstrom, William: Bingley: Emerald Publishing Limited.

Kremer, M., J. Willis, and Y. You (2021). Converging to convergence. *National Bureau of Economic Research*.

Mitchener, K. and I. McLean (1999). U.s. regional growth and convergence, 1880-1980. *Journal of Economic History 59*(4), 1016–1042.

Rodrik, D. (2013). Unconditional convergence in manufacturing. *The Quarterly Journal of Economics 128*(1), 165–204.

Rosenbloom, J. (2002). *Looking for Work, Searching for Workers: American Labor Markets during Industrialization*. New York: Cambridge University Press.

Steven Ruggles, Sarah Flood, S. F. R. G. J. P. M. S. and M. Sobek. (2021). Ipums usa: Version 11.0 [dataset]. minneapolis, mn: Ipums.

Wright, G. (1999). *Old South, New South*. New York: Basic Books.

Young, A., M. Higgins, and D. Levy (2008). Sigma convergence versus beta convergence: Evidence from u.s. county-level data. *Journal of Money, Credit and Banking 40*(5), 1083–1093.

5 Appendix

Table A1: Regional Price Deflators by Division and Year

Division	1900	1910	1920	1930	1940
Pacific	90.18	95.02	100.72	98.88	96.95
Mountain	90.65	94.14	99.00	98.40	97.70
West North Central	107.29	104.40	101.16	100.60	100.00
East North Central	105.66	102.79	100.00	100.00	100.00
West South Central	96.72	97.74	98.61	99.27	100.00
East South Central	106.12	101.98	98.08	100.00	102.04
South Atlantic	98.11	98.31	98.46	98.54	98.61
Middle Atlantic	98.59	100.72	102.96	101.50	100.00
New England	96.24	95.72	95.16	96.43	97.66

Note: This table lists the regional price deflators used in the primary analysis. These values come from Table 1 of Mitchener and McLean (1999). Values for 1910 and 1930 were linearly imputed. The regional deflators for 1940 were used for 1947 as well.

Table A2: Industry Group Names

20	Food and Kindred Products	21	Tobacco Products
22	Textile Mill Products	23	Apparel and Other Textile Products
24	Lumber and Wood Products	25	Furniture and Fixtures
26	Paper and Allied Products	27	Printing and Publishing
28	Chemical and Allied Products	29	Petroleum and Coal Products
30	Rubber and Plastics Products	31	Leather and Leather Products
32	Stone, Clay, and Glass Products	33	Primary Metal Industries
34	Fabricated Metal Products	35	Industrial Machinery
36	Electronics	37	Transportation Equipment
38	Instruments (Technical)	39	Miscellaneous

Note: This table lists the twenty SIC 2 industry groups