Still Growing Together? The Spatial Distribution and Industrial Composition of U.S. County GDP since 1870*

Scott L. Fulford† Fabio Schiantarelli‡

March 2026

Abstract

We construct the first estimates of U.S. county GDP by broadly defined industrial sectors from 1870 to 2018. GDP per worker inequality has two broad eras: from 1870 to 1970 U.S. counties grew together, as states converged and inequality within states fell. Manufacturing became less concentrated and more equally productive, explaining much of inequality's decline. But U.S. counties stopped growing together after 1970. GDP inequality within states increased as the most productive metro areas decoupled from the rest of the state. Replacing manufacturing as the path to riches, tradable services, such as finance and business services, became increasingly concentrated in rich areas, while poor areas' economies became dominated by government, education, and health services. Troublingly for growth, the most productive counties' populations no longer grow more rapidly. We show county productivity convergence and the subsequent divergence had a substantial impact on growth since 1870.

JEL classification: R12; O4; N9
Keywords: Regional Growth; Inequality; Productivity; Industrial Composition; County growth; United States; GDP

Preliminary, please do not quote without consulting the authors.

*This paper has benefited from seminars at Harvard University, Vanderbilt University, University of California, Merced, and West Virginia University, presentations at the 2025 NBER Summer Institute Development of the American Economy conference, the 2020 American Real Estate and Urban Economics Association Meetings and the Sembenelli Day Conference at the University of Turin, as well as conversations with and suggestions from Susanto Basu, Klaus Desmet, Oded Galor, Bob Margo, Ivan Petkov, and David Weil. We also benefited from the detailed comments by Andrea Cerrato and from the research assistance by Liyang Hong, who also contributed to the derivation of the sectoral decomposition of the Theil index. We are, finally, grateful to Edoardo Bella, Enrico Cristoforoni, and Matteo Masullo for their careful reading of the paper and for their suggestions and to Matteo Masullo for his research assistantship. The authors have no conflicts of interest to declare.

†Consumer Financial Protection Bureau, scott.fulford@cfpb.gov. The views expressed here are those of the authors' and do not necessarily represent the views of the Consumer Financial Protection Bureau or the United States. This work was completed without using government resources.

‡Boston College Department of Economics and IZA, schianta@bc.edu. 140 Commonwealth Ave, Chestnut Hill, MA 02467.

1 Introduction

Atlanta, Georgia, is the heart of the "New South." When Henry W. Grady, the editor of the Atlanta Constitution, coined the term in 1874, the GDP per worker of what would become the Atlanta metropolitan area was one quarter of New York City's and half of Savannah, Georgia's, then a major port city. Rural Georgia was about half as productive as rural New York at the time, with its prosperous upstate farms. Yet Atlanta grew. Coca Cola was founded there in 1886. Delta Airlines moved its headquarters there in 1941. Atlanta's GDP per worker surpassed Savannah's around 1960 and Rochester, New York's around 1990. Atlanta was nearly as productive as New York City in 2000. At the same time, the economies of rural Georgia and rural New York converged, as did Savannah's and Rochester's, so that the overall GDP per worker of New York State and Georgia had largely converged by 1970. Yet since then, Atlanta and New York City's GDP per worker have grown faster than their states' rural and smaller metropolitan areas.

What do these different growth paths tell us? We examine how U.S. counties' gross domestic product (GDP) has evolved from 1870 to 2018. Understanding how local economies grow together or apart, with their borders open to labor, capital, and trade gives insight into fundamental questions about growth, development, regional inequalities, and economic history. Further, the U.S. and other countries devote substantial resources to location-specific policies to help develop local economies (Austin, Glaeser, and Summers, 2018). These policies' effectiveness rest in part on understanding how local economies have developed across time and industries. As recent research on the relationship between trade and manufacturing illustrates (Autor, Dorn, and Hanson, 2016), a positive aggregate economic trend can have devastating local economic effects and cause important political shifts (Autor et al., 2020).

We build a new data set that measures the production that takes place within each county by broad industrial sector. The data set allows for an industrial breakdown between agriculture, manufacturing, mining, construction, and different types of services, and makes it possible to analyze local economic development and its industrial composition over a long period at a level of spatial disaggregation unavailable up to now. Our approach, which builds on Fulford, Petkov, and Schiantarelli (2020), uses earnings and employment information contained in individual census records or in county employment records to reconstruct where economic activity occurred in services. We also use direct measures of output and inputs when they are available for manufacturing and agriculture. We thus contribute to a growing literature using innovative sources, such as satellite light density (Henderson, Storeygard, and Weil, 2012), to infer local economic activity. We justify our approach using standard economic theory for allocating GDP using the wage bill. We show our estimates closely compare to the recent BEA county GDP estimates after 2001. Yet our estimates are the only ones available for county GDP before 2001 and, aggregated up to states, the only

estimates of state GDP before 1963, to our knowledge. In this sense, we follow in the footsteps of the efforts to produce the first National Income and Product Accounts at the National Bureau of Economic Research and what would become the Bureau of Economic Analysis within the Bureau of Commerce.

Although there is more nuance, our new data tell a tale of two eras: from 1870 to 1970, U.S. counties grew together, as states converged and inequality within states fell. Manufacturing became less concentrated and more equally productive, explaining much of inequality's decline. This convergence explains a large portion of aggregate growth from 1870 to 1970; much of growth during the period was regional "catch up" growth as Atlanta approached New York City. But U.S. counties stopped growing together after 1970. Inequality within states increased as the most productive metro areas increasingly decoupled from the rest of the state, as the experience of New York and Georgia illustrates. GDP per person inequality increased even more than GDP per worker, driven by increasing employment-rate inequality. Replacing manufacturing as the path to riches, tradable services, such as finance and business services, became increasingly concentrated in rich areas, while poor areas' economies became dominated by government, education, and health services. Many of the forces pushing for convergence ceased and the most productive counties' populations no longer grow more rapidly than other counties. Growth since 1970 would have been faster had inequality not increased.

We structure our evidence by highlighting a series of key results that characterize the distribution of county-level economic development since 1870. We illustrate the changes through the experiences of New York and Georgia, but similar stories describe many other regions such as the diverging experiences of Boston and Springfield in Western Massachusetts.

> **Result 1**: U.S. counties tended to converge from 1870 until 1970 but subsequently grew apart.

We focus on the Theil inequality index because it is additively decomposable (Theil, 1967; Cowell, 2000), however, this result is robust to using other inequality measures. The convergence and subsequent divergence also occur when we limit the sample to metropolitan areas or to non-metropolitan areas.

> **Result 2**: Falling inequality between states explains most of the fall in nationwide county inequality from 1870 to 1970. After 1970, increasing inequality within states explains most of the overall inequality increase.

While in the late nineteenth century, the New York state and Georgia's productivity was very different, they converged until 1970 as Atlanta approached the GDP per worker of New York City. After that, Atlanta and New York City both diverged from their hinterlands and secondary metro areas. Rochester, NY had an identical GDP per worker to New York City from 1910 to 1970 but

grew less quickly than New York City after that. Rochester has instead converged with Savannah, Georgia, which Atlanta passed in GDP per worker in 1960.

> **Result 3**: Since 1970, GDP per person inequality increased more than GDP per worker inequality. The difference is due to a growing inequality in the employment-to-population ratio not due to changes in the prime-age population share.

With the growth of retirement-oriented states such as Florida or Arizona, one hypothesis is that growing inequality in the prime-age share is responsible for the difference between GDP per worker and GDP per person inequality. Instead, we show that inequality in the share of the population aged 25 to 54 appears constant since 1950, while growing employment-to-population ratio inequality—where some counties have far more people out of the labor force or unemployed—appears responsible for the difference.

> **Result 4**: Before 1970 there is a negative relationship between state GDP per worker and inequality; more productive states were more equal. The relationship becomes positive in the 1970s; more productive states were more unequal, particularly in recent years.

New York State, for example, had a fairly equal distribution of GDP per worker across counties from 1870 to 1970 characterized by productive upstate farms and smaller industrial cities such as Rochester, while Georgia was markedly more unequal and much poorer. In more recent decades, this relationship has flipped as the most productive metro areas in the richest states pull away from both rural areas and other cities.

Combining Results 1 through 4, we show that U.S. geographic inequality is no longer primarily about differences between regions or states, but instead about differences within them. This result complicates and enriches discussions about convergence (see Barro and Sala-i-Martin (1991) and Barro and Sala-i-Martin (2004) for a discussion of state convergence). On the one hand, U.S. states have converged over the long term. On the other hand, convergence has largely stopped, and there are growing inequalities within states, particularly in the highest GDP per worker states.

We next examine the different sectors' contribution to the evolution of county GDP per worker inequality. Since 1870, there have been four important sectoral transitions: (1) the declining importance of agriculture; (2) the increasing and then decreasing importance of manufacturing; (3) the slow increase in education and health services over the entire period; and (4) the more recent rise of what we refer to as "tradable" services such as finance, insurance, professional, scientific, and business services. These services are tradable in the sense that consumption and production can be geographically separated, unlike personal or recreational services, which generally must be consumed where they are produced. Our data uniquely capture services' importance and distribution which is crucial for understanding sectoral transformations because services have been more than 60 percent of GDP since 1880.

To understand what role different sectors play, we build on Shorrocks (1982) and decompose the Theil index into components that show the contribution of each sector. Our decomposition shows how two components help to determine an industry's inequality contribution: (1) its overall share, and (2) the correlation across counties between the industry share and county GDP per worker. Industries concentrated in rich counties will tend to increase inequality. Four major facts stand out:

> **Result 5:** Agriculture used to be the primary activity of the poorest counties. Now, the poorer the county, the larger the share in government, education, and health.
>
> **Result 6:** The path to riches has changed from manufacturing to tradable services. From 1870 to 1950 manufacturing became increasingly concentrated in the richest counties. The manufacturing share is now the highest in middle income counties, while the richest counties increasingly produce tradable services.
>
> **Result 7:** Manufacturing's contribution to inequality is the largest before 1960 and its decline is the main explanation for the fall in inequality from 1930 to 1970. Manufacturing's declining contribution to inequality is explained by both greater geographical diffusion and falling cross-county productivity inequality. Finance and professional services contribute to the rise in GDP per worker inequality, mostly because they have become more concentrated.
>
> **Result 8:** Before 1950, government services increased inequality, but after 1950, government services reduce inequality.

Our results 5 through 8 describe the complex interaction between sectoral transitions and geographical inequality. Some changes—such as the spreading out and productivity homogenization of manufacturing—have reduced inequality. Others—such as the increasing concentration of tradable services—increase it. Meanwhile, government services reduce inequality directly and government transfers reduce inequality indirectly by supporting local non-tradable services such as health and education that have come to dominate poor counties' economies.

While inequality is important on its own, convergence also explain a substantial portion of overall U.S. growth. To capture the importance of changing inequality, we introduce a simple Productivity Gap Index which measures how much higher GDP would be if, hypothetically, people in low productivity counties could be brought up to the productivity of high productivity counties either by moving, trading, or using better technology. It thus measures the combination of housing (Hsieh and Moretti, 2019), transportation (Donaldson and Hornbeck, 2016), and other geographic missallocations or frictions (Hsieh and Klenow, 2009, 2010) that reduce productivity. Unsurprisingly, given the amount of geographic inequality we find, the Productivity Gap Index is large, but more important, it changes over time. Using one measure of frontier productivity, we calculate that 39 percent of real GDP per worker growth between 1870 and 1970 was from the convergence in productivity as laggard counties caught up. This calculation is sensitive to the frontier defini-

tion but the importance of convergence remains. But as convergence slowed, so too did catchup growth, and it went into reverse after 1970. If the productivity gap had held constant at the 1970 level, rather than increasing, annual GDP per worker growth would have been 22 percent faster between 1970 and 2018.

We examine several changing factors that adversely affect convergence: declining population mobility to where it has the highest returns, decreasing education convergence, and an increase in agglomeration effects. We show that while the population of the highest GDP per worker counties used to grow much faster than other counties, they no longer grown much more rapidly. Indeed, during the 1980s and 1990s high GDP per worker counties' population grew *slower* than in low GDP per worker counties. A decline in factor mobility to where it has the highest returns is a troubling development for long-term growth. Cross-country education inequality declined substantially from 1870 to 1970. Since then, education convergence has slowed, and the top metropolitan statistical areas (MSAs) have a growing advantage in attracting college educated workers. Finally, tradable services and manufacturing productivity in the MSAs with the highest GDP increased more quickly than in other MSAs and non-metro areas since 1970, suggesting growing agglomeration economies. We summarize these trends as:

> **Result 9:** Population growth and education used to be strongly pro-convergence, but after 1970 became neutral or anti-convergence. At the same time, agglomeration effects in manufacturing and tradable services appear to have increased.

Relationship with the literature. Our work touches on several research areas. First, there is the large literature on convergence within the U.S. (Barro and Sala-i-Martin, 1991, 1992, 2004; Barro, 2015) that typically is very supportive of relative (or "β") convergence, using 1880–2000 state level data on personal income per person.[1] There is similar evidence in favor of relative convergence for regions in Europe and prefectures in Japan. However, while absolute (or "σ") convergence for U.S. states was rapid until around 1960, it was approximately zero after. Using county level data for personal income per person for the period 1970-1998, Young, Higgins, and Levy (2008) show that there is evidence of significant absolute divergence for that period, but this is not true for county income inequality measured by the Gini coefficient. Higgins, Levy, and Young (2006) present evidence supporting relative convergence for U.S. counties during the same period using personal income per person.

Second, a more recent literature is mostly concerned with the regional, state, and city divergence after 1970 and its reasons. Moretti (2012), Hsieh and Moretti 2019, and Eckert (2019) emphasize the large and growing wage dispersion across U.S. cities since the 1960s. Hsieh and Moretti (2019) use an spatial equilibrium model to estimate that the lack of population mobility

[1] See also Blanchard and Katz (1992) who discuss the general features of regional booms and slumps studying the behavior of U.S. state level employment, unemployment and wages over the last 40 years.

to the most productive cities has large economic costs. Ganong and Shoag (2017) focus on states and examine the importance of state housing regulations in slowing this mobility. A recent literature has focused on the local effects of increased manufacturing competition from China (Autor, Dorn, and Hanson, 2016, 2021). Giannone (2019) focuses, instead, on the structural reasons for diverging productivity and analyzes the role of skill-biased technological change and agglomeration economies for wage convergence between 1980 and 2010. Kleinman, Liu, and Redding (2023) use a dynamic spatial general equilibrium model to study how the interaction between capital and labor explains the decline in income convergence across states since 1965. The importance of both agglomeration economies and housing supply factors in models of spatial equilibrium is emphasized and reviewed by Glaeser and Gottlieb (2009). Gennaioli et al. (2013) instead focuses on role of human capital in regional development. This literature relies mostly on wages or income per capita as an indicator of productivity dispersion and is mostly at the metropolitan level, while we rely on direct measures of labor productivity (real GDP per worker) in each sector at the county level calculated over a longer period of time. Our more granular data allow us to distinguish the role of inter-state and intra-state components in accounting for changes in inequality in a broad historical context. Finally, our data allows us to describe how the process of structural transformation interacts with the spatial evolution of income. A related literature focuses on income and wage inequality across people rather than its spatial component.[2] Our granular data show that inequality has a strong spatial component which contributes to and complicates inequality across people.

Third, the emphasis on GDP per worker and its sectoral breakdown is shared by several empirical and theoretical contributions. On the empirical front, Bernard and Jones (1996) uses state level GDP data per worker and its sectoral breakdown for the period 1963–1989. They find evidence of absolute and relative convergence at the aggregate state level until the early 1980s, while sectors behave heterogeneously, with only manufacturing and mining displaying convergence in labor productivity. Aggregate convergence is mostly accounted for by manufacturing's productivity convergence. We also find a decrease in manufacturing's dispersion and productivity inequality across counties during the period.

More generally, a broad literature uses theoretical models to understand overall growth through sectoral growth (see Herrendorf, Rogerson, and Valentinyi (2014) and Breinlich, Ottaviano, and Temple (2014) for reviews). Related theoretical and empirical work emphasizes the role of structural transformation for spatial growth and the process of regional convergence (Caselli and Coleman, 2001; Michaels, Rauch, and Redding, 2012; Comin, Lashkari, and Mestieri, 2021). Eckert and Peters (2022) model technological catch-up and its interaction with geographic specialization

[2]See, for instance Goldin and Margo (1992) who focus on the compression of wages in the mid 20th century, Lindert and Williamson (2016) and Piketty and Saez (2003) on decrease and then increase in income inequality in the 20th century, and Goldin and Katz (2010) on the role of education an technology in the evolution of wage differentials. Finally, a recent literature examines the geography of intergenerational mobility (Chetty et al., 2014).

with an application to U.S rural areas from 1880 to 1920. Desmet and Rossi-Hansberg (2009) and Desmet and Rossi-Hansberg (2014) focus on the transition between manufacturing and services and produce models that help explaining some of the stylized facts in the spatial growth distribution and its sectoral composition. Our granular data covering a long period gives us a deeper understanding of sectoral transformation's role in affecting the county-level distribution of productivity. We document the relative role of concentration and within industry productivity dispersion and, most importantly, we are the first to quantify the contribution of each sector to the dispersion of GDP per worker across counties.

Our finding that some services such as personal and government services push towards convergence, while tradable services push towards divergence is similar to the observation by Eichengreen and Gupta (2013) on the two waves of service sector growth and Desmet and Rossi-Hansberg (2014) on the importance of the spatial concentration of services. Again, our finer spatial disaggregation at the county level, and the much longer time period covered is a distinctive feature of our contribution that allows us to distinguish the role of concentration and productivity as a source of GDP per worker inequality.

2 Construction of county GDP

This section describes our construction of county level Gross Domestic Product (GDP) at current prices from 1870 to 2018.[3] The basic issue in obtaining a measure of sectoral GDP at the county level over a long period of time is the limited information on output and inputs by sectors in the past. The value of both output and inputs is available only for manufacturing and the value of output only for agriculture. For all the other industries, including services, mining, and construction, one must resort to an indirect method. Our main approach is to build on Fulford, Petkov, and Schiantarelli (2020) and use relative wage earnings or employment by county to allocate value added at the state level for each sector to individual counties. The details of our calculations vary from year to year based on the available data. We describe the broad outlines of our calculations below, but the full details are in Appendix B.

We extend Fulford, Petkov, and Schiantarelli (2020, FPS) in several ways. First, we create consistently measured individual sectors. Second, FPS used county-level personal income to project forward county level GDP after 1950, in this paper we build a direct GDP measure using the same basic methodology over the entire period. Finally, we produce county level GDP by industry, while FPS calculated total GDP at the county-group level to match their ancestry data.

[3] In principle, the data exist to apply our approach in 1850 and 1860. However, the massive structural change of the Civil War, particularly manumission, and the substantial expansion in the number of counties as population shifted west suggest that 1850 and 1860 would not be comparable to later periods.

We calculate county GDP for the sixteen industries for which we can construct consistent measures over time and across space: 1) Agriculture; 2) Mining; 3) Construction; 4) Manufacturing;[4] 5) Transportation; 6) Communication; 7) Public utilities; 8) Wholesale trade; 9) Retail trade; 10) Finance and Insurance; 11) Real Estate; 12) Professional services (Professional, Scientific, Technical, and Business); 13) Education and Health services (Education, Health, and Social Services); 14) Recreation services (Recreation, Arts, Entertainment, and Accommodation); 15) Personal services (Personal, Domestic, Repair, and Other);[5] and 16) Government services.

2.1 Estimating GDP when not directly available

In this subsection, we outline the assumptions about production and markets that would justify our indirect approach of allocating GDP to a county according to relative wage earnings. Consider a sector such as transportation services in state s. We omit the sector subscript for ease of notation. The nominal state GDP in this sector is the sum of each county's (c) GDP:

$$Y_s = \sum_{c=1}^{C} Y_c. \tag{1}$$

Suppose that within each county c, the production by a representative perfectly competitive firm is characterized by a Cobb-Douglas production function and that the sectoral prices are common across all counties. Then nominal GDP is:

$$Y_c = PA_c K_c^{\alpha} L_c^{1-\alpha}.$$

Thus, if labor is paid its marginal revenue product, wages in a county must be:

$$w_c = MRPL_c = P(1-\alpha)A_c K_c^{\alpha} L_c^{-\alpha} = (1-\alpha)Y_c/L_c. \tag{2}$$

Combining equations (1) and (2), county level production is just the fraction of the state wage bill in that county:

$$Y_c = \frac{w_c L_c}{\sum_{c \in S} w_c L_c} Y_s. \tag{3}$$

[4]We can only make the division between durable and non-durable manufacturing after 1963 and have found that even then the division is not possible for smaller counties due to data not being made public to prevent disclosure (there may only be one durable goods manufacturer in a county), and so we do not break down manufacturing between durables and non durables, or by other categories, such as "high-tech" manufacturing industries, even if such a division is possible in some periods.

[5]The breakdown of Services into the four components of Professional, Education and Health, Recreation, and Personal is driven by the ability to allocate employment or earnings into these broader groupings consistently from 1870. Narrower categories are possible in some years, but not consistently across the period.

This approach allocates state GDP to counties based on where the labor earnings in that sector occur. While this approach assumes the labor share is the same for all counties within a state for a given sector, it does not assume a common labor share across industries or states. For exposition, we develop Equation 3 assuming firms are perfectly competitive, but we can easily generalize to the case of imperfect competition in the output market. Provided there is a common markup for all firms within a sector in a given state, Equation 3 would still hold.

2.2 Sources and methods by period

We will use Equation 3, or variations of it with stronger assumptions depending on the available data, to construct county GDP. Based on available data, we calculate county GDP at a decadal frequency up to 1950, in 1958 and 1963, and at an annual frequency from 1969. Our calculations require two steps, with some variation by period: (1) We allocate national value added by industry to states; and (2) we allocate state value added to counties.

State GDP by industry

Our first step is to calculate state GDP for the entire period. More details are contained in Appendix B.1. Starting in 1963, the BEA produced yearly estimates of state GDP by industry (U.S. Bureau of Economic Analysis, 2016), but no estimates of industry-state GDP exist before then. We project state GDP by industry backwards from 1963 to 1958 using the national industry growth rate to get estimates of state GDP for 1958. Before 1958, we rely on Equation 3 at the state level to allocate national value added in most sectors, except manufacturing, for which we have direct estimates of output and inputs, and agriculture, for which we have information on output and construct an estimate of inputs using the national input-to-output ratio.

To construct the state wage bill before 1958, we first collect estimates of state wages by industry then combine them with state employment. The 1940 census asked respondents about wages and income. We use the full-count individual records from the 1940 census to construct the average wage within each state and industry and assume that the same relative wage applies in 1930 and 1950. Before 1930, we estimate the wage distribution across states from a variety of historical sources. For example, we create a government services wage by taking an average of postal workers, municipal laborers in sanitation and sewage, and police detectives. Similarly, we use an average of two horse teamster and locomotive engineer wages to create a transportation wage distribution by state. Appendix B.1 details our construction and sources for each industry. The necessary assumption is that the wages we collect reflect the cross-state wage distribution in that industry, not that they are the average wage in the industry.

We combine our state wage by industry estimates with state employment from the decadal

census individual records by allocating a census occupation to each industry (Bureau of the Census, 1950) to create state wage bill estimates. We use the wage bill to allocate national value added by industry to states. The National Income and Product Accounts (United States Department of Commerce, 1993) breaks down national value added by industry starting in 1929. Before that we use Gallman and Weiss (1969) for services and Wright (2006) for mining.

County GDP by industry

Our second step is to allocate state industry GDP to counties using Equation 3. The county-level information varies by period, so we summarize the difference by period below.[6]

County GDP in 1969 and after. In 1969, the BLS started publishing yearly county employment compensation by industry, so from 1969 we have all of the information required to construct the wage bill and to allocate state industry GDP (available since 1963) to each county using Equation 3.

County GDP in 1930 and 1940. The 1940 census asked respondents about wages and income. We use the full-count individual records from the 1940 census to construct the wage bill and the average wage within each county and industry and then apply Equation 3. We assume the same geographic distribution of relative wages within each industry holds in 1930 but use the 1930 census to calculate industry employment to create the county wage bill.

County GDP in 1950, 1958 and 1963. We use the City and County Databooks (United States Department of Commerce Bureau of the Census, 2012) for 1950, 1958, and 1963. Our approach is a hybrid of our approach before 1930 and after 1969 depending on whether employment, earnings, or value added is available for each industry.

County GDP from 1870 to 1920. To allocate state GDP before 1930, we use the relative wages in urban and rural areas in the 1940 census to create an estimate of the state urban-rural wage distribution by industry before 1930. We assume the relative urban-rural wages in 1940 holds before that. We then use available wages to calculate the wage bill for urban and rural areas.[7]

[6]In the census data, employment and earnings are based on place of residency and not on place of work, so commuting across county lines will overestimate GDP in some areas and underestimate it others for industries other than agriculture and manufacturing. Starting in 1969, we use BEA county earnings which are allocated by place of work, so commuting does not affect our estimates during the period when cross-county commutes would have been the largest problem.We discuss the sources and location of residence and jobs in more detail in Appendix Section B.2.

[7]The exact procedure depends on what kind of wage we observe, since sometimes we observe a state average and sometimes only an urban wage. For example, we use the state average teacher salary for wages in the education sector. To find the urban and rural wage bill in the years before 1930, we use education employment in urban and rural areas,

Given employment in education in the state for urban and rural areas from the census, we can obtain the respective wage bill. We then allocate state GDP by industry to these urban and rural groupings in proportion to their wage bill. Finally, we assume that within each state-industry-urban/rural grouping, wages are the same so we can use employment to allocate each industry GDP to counties in the state. If wages are equalized across counties in a state within an urban (rural) location, then wages cancel out in Equation 3 and county-industry GDP for urban counties is:

$$Y_c = \left(\frac{L_c}{\sum_{c \in S,Urban} L_c)} \right) Y_{S,Urban}. \tag{4}$$

This final step improves on Fulford, Petkov, and Schiantarelli (2020) who allocated national value added in a sector using employment, assuming that wages are equalized nationally in each sector. By constructing a state and industry wage distribution and a relative urban-rural wage distribution by state and industry, we can relax this assumption and only assume that in each sector wages are equalized within urban or rural areas within a state.

Finance and real estate

Finance, insurance, and real estate pose special challenges. For finance and insurance, we use state banking capital, rather than employment and wages, to allocate national finance value added to states before 1929 because the state distribution for finance wages is unreliable before 1930, while banking capital is better measured at the state level. Using the marginal revenue product of capital instead of labor produces equations similar to Equations 2 and 3 (see Appendix B for derivations and data sources). If the risk adjusted returns are equalized, which as has some justification as shown by James (1976), we can then use financial sector capital to allocate national GDP in finance to each state. The allocation of state value added to each county is then based on the relative wage bill or employment within each state.

Real estate poses a special problem which we largely sidestep. Existing buildings produce a flow of services which National Income and Product accounting considers part of real estate in GDP. These flows are typically measured as rent and owner imputed rent. These service flows are important because they allow the NIPA expenditure, income, and production concepts to match. We capture the construction of new buildings, but measuring the continuing service value of existing structures is harder. Moreover, while rental and leasing employment is part of real estate, employment in real estate is not a good way to allocate these existing structure service flows, which make up most of real estate GDP. Instead of using the wage bill to allocate real estate value added to counties, we instead allocate it proportionally based on county GDP without real estate. This

and solve for the relative urban and rural wage that delivers the urban-rural wage bill ratio observed in 1940 that is consistent with the state average teacher wages for that year.

approach means that although our measure includes real estate for consistency with existing NIPA practices, real estate is not a direct contributor to county inequality either positively or negatively.

Consistent counties and constructing industry employment

Several other data construction elements support our analysis. Appendix B.2 explains how we construct industry employment from the individual census records. Appendix B.3 discusses how we construct consistent counties over the entire period. There are relatively few changes, but counties do occasionally split and combine and keeping track of them over a long period requires care. Finally, Appendix B.4 describes the construction of county income per person during the sub-period 1950-2018 when it is available for comparison.

2.3 Comparison of our county GDP measure to other approaches

Because ours are the first estimates of county GDP, our estimates at the county level have no direct comparison for earlier periods. However, combing our county GDP into states or nationally, we produces GDP estimates that are comparable to other estimates. We show these comparisons in detail in Appendix C. Our national GDP estimate constructed by combining industry value added compares closely to other estimates of national GDP. Before 1969, no state GDP estimates exist, so the sum of our counties is the first estimates of state GDP for much of the period, but we compare our state estimates to early estimates of state personal income per person which have been used to study state convergence. Although the concepts are slightly different, our measure closely approximates personal income per person.

Figure A-12 in the Appendix shows that our approach matches sectoral shares of the economy provided by the statistical agencies as well, when these become available after 1929, and provide an independent measure of the industrial composition before that. Indeed, given the difficulty in calculating shares starting in 1929 and during the subsequent Depression, there is some reason to prefer our more stable estimates in the early years.

Finally, the BEA started calculating county GDP using non-public information in 2019 for the period since 2001 and we compare our county estimates to the BEA's estimates. Appendix C shows that excluding counties where mining is more than 40 percent of the economy or utilities are more than 25 percent, our measure lines up nearly exactly with the BEA measure. Excluding these counties, over the period when our measure and the one provided by the BEA overlap, the slope coefficient regressing our log GDP per person on log BEA GDP per person is statistically equal to one, the constant is zero, and the R^2 is 0.92. Because our measure closely corresponds to a county GDP estimates after 2001 calculated by the BEA, these results suggest it is likely to be accurate before 2001 as well.

Main sample. For our main sample, we exclude these heavily mining or utilities counties. Our employment or wage-based allocation of GDP to counties may not be as accurate for mining, so we do not want our measurement of inequality to be affected by including these high-mining counties. In excluding some counties, we divide our total period into long sub-periods to maintain as constant a sample as possible while allowing counties to shift mining and utility concentration. From 1980 on, we exclude counties whose average mining share in the BEA estimates from 2001–2019 was greater than 40 percent and whose utility share was greater than 25 percent. From 1940 to 1979, we exclude counties whose average share mining during the same period was greater than 40 percent and utilities greater than 25 percent. We do the same exclusion for the period 1870 to 1930. In all years, we exclude counties with a population less than 2500. Together, these restrictions exclude from our main sample 6.5 percent of county-years for which we can calculate GDP. Appendix Figure A-1 shows the inequality measure we discuss in the next section with and without the sample restriction. The inequality measure is nearly the same and our exclusions tend to limit inequality's rise.

3 The distribution of economic activity over space and time

This section examines how the distribution of GDP per worker and per person has evolved over space and time. In the next section, we will discuss how these changes are related to the changes in industrial composition at the local level. Where appropriate, we use a GDP deflator from Sutch (2006) updated using the chained BEA GDP deflator. Our core analysis of inequality is scale invariant, so the use or not of a common deflator does not affect our conclusions.

3.1 Inequality over time

We use the Theil inequality index to describe the evolution of GDP per worker:

$$T = \frac{1}{n} \sum_{c} \left(\frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}} \right)$$

where y_c is GDP per worker whose mean is $\bar{y} = \frac{1}{n} \sum_c y_c$. Unlike many other measures, the Theil index is additively separable, so that we can decompose overall inequality into an across-state and a within-state component. The Theil index ranges from 0 representing no inequality (when $y_c = \bar{y}$) to indefinitely large as GDP per worker is concentrated in a single county. Higher values represent more inequality.

Figure 1 shows GDP per worker inequality over time. The shaded areas divide inequality into within state inequality and between state inequality. The Theil index is scale invariant so it does

not matter whether GDP per worker is deflated by a common deflator or not.

> **Result 1**: U.S. counties tended to converge from 1870 until 1970 but subsequently grew apart.

Our conclusions are not affect by choosing a specific inequality index. Appendix Figure A-2 shows the Gini coefficient and the standard deviation of log GDP per worker. All inequality measures show a substantial shrinking of the distribution until about 1970 followed by an increase.

Figure 1 also reveals that most of the inequality decline until 1970 is between states, although within-state inequality also declines. The increase since 1970 has been mostly within states. There is almost no increase in across state GDP per worker inequality. We summarize this contrasting pattern as:

> **Result 2**: Falling inequality between states explains most of the fall in county inequality from 1870 to 1970. After 1970, increasing inequality within states explains most of the overall inequality increase.

Our Results 1 and 2 contribute to and extend a large literature (Barro and Sala-i-Martin, 1991, 1992, 2004) that has found that U.S. states converged both relatively (β convergence) and absolutely (σ convergence) from 1880 until 1960 using state level data on per capita personal income and that state convergence has slowed or stopped since then. The century-long inequality decrease starting in 1870 is likely connected to several factors including labor's movement out of agriculture and into manufacturing and services (Caselli and Coleman, 2001; Eckert and Peters, 2022; Desmet and Rossi-Hansberg, 2009; Michaels, Rauch, and Redding, 2012; Desmet and Rossi-Hansberg, 2014), the associated migrations out of the rural South towards cities in the North-East, Mid-West, and the West particularly associated with the Great Migration, and to increased economic integration through improved transportation (Atack, Margo, and Rhode, 2022). While state-level convergence was the primary contributory factor early on, our county measures show that state convergence has mostly stopped and inequality is now explained mostly by within state inequality. The spatial granularity of our county-level data distinguishes our contribution and allows us to decompose the evolution of GDP per worker inequality in the U.S. into an across-state and a within-state component.

Figure 2 Panel (A) shows GDP per person inequality, which is higher than GDP per worker inequality but fell at a similar rate until 1970. Since then, however, GDP per worker inequality has increased much more rapidly. The difference between GDP per person and GDP per worker is the employment to population ratio. The gap between per person and per worker inequality has grown since 1970, suggesting that employment inequality has increased.

Panel (B) shows two possible explanations for the difference between GDP per worker and GDP per person inequality. . The employment-topopulation ratio is the total employed in a county

compared to the total population. The sharp increase in employment-to-population inequality between 1950 and 1969 is partly driven by a definition change driven by available data. In 1950 and earlier, we measure the employment-to-population by the share of people with a market occupation in the census micro-data. In 1969 and after, we measure employment-to-population by the number of jobs from the BLS divided by the population (between 1950 and 1969, the definition varies by industry). The prime-age share is the share of the population aged 25-54. While prime-age share inequality has been relatively flat since 1970, employment-to-population inequality has been increasing. Panel (B) suggests the growth of new retirement communities in places like Florida and Arizona and the "hollowing out" of rural areas as young people leave, are not major contributors to county inequality. Demographic change's overall contribution to inequality changes since 1970 appear minimal; instead, the difference between GDP per person and per worker appears to come from employment. We summarize these results as:

> **Result 3**: Since 1970, GDP per person inequality increased more than GDP per worker inequality. The difference is due to a growing inequality in the employment-to-population ratio not due to changes in the prime-age population share.

3.2 The geography of GDP per worker

The rise in within state GDP per worker inequality changed the geography of production in the U.S. Figures 3 and 4 map the county difference from overall U.S. GDP per worker in 1880, 1920, 1970, and 2018. Counties that are richer than the average appear in blue, counties that are poorer in red, and counties about average are in yellow. Production tends to be concentrated in highly populated, but geographically small, metropolitan areas, so below average (red) counties take up more surface area than the above average (blue) counties. Put a different way, more than half of counties are below U.S. GDP per capita and these counties tend to be large in area.

In 1880 and 1920 there are clear regional differences. The Northeast and Midwest are blue and yellow, as are California, Washington and Oregon. While there are pockets of blue or yellow in the South, most Southern counties were much poorer than average. By 1970 this pattern had shifted, and regional differences were much less clear. By 2018 it had completely shifted; there were no longer fairly uniformly richer regions and poorer regions. The entire country had become a patchwork of small dots of blue and yellow, surrounded by a sea of red. The maps show the geographic effect of the declining across-state inequality in Figures 1.

There has been a corresponding shift in the production center of the U.S. that helps show that regional convergence is largely finished. The square in each map is the geographic center production in the U.S. (using county centroids weighted by county GDP) and the circle is the geographic population center. In 1880, following the Civil War and incomplete Reconstruction,

the square is far north, in a line with Philadelphia and just south of Detroit. Over the next 130 years the square grows and slowly shifts south and west. In 1880, the circle is much further south than the square, showing that while production is mostly in the Northeast and Midwest, the population is more evenly spread. By 1970, they nearly coincide as the square moves west and south as the relative GDP of these regions grows and the center of the U.S. economically was also its population center in 2018.

The differing stories of Atlanta and the rest of Georgia and New York City and upstate New York illustrate the new geography of inequality. Atlanta has been growing quickly, and while not as productive as New York City, it has converged substantially. It has brought its closely surrounding counties along, but it has diverged from the rural areas of Georgia and secondary cities. Appendix Figure A-3 shows the GDP per worker of New York City and Rochester, New York, Atlanta and Savannah, Georgia, and both states non-metropolitan areas. The importance and size of the economy of Atlanta means that the state of Georgia has nearly converged to New York state, but the large gains are mostly for Atlanta and adjacent areas. Secondary Georgia cities, such as Savannah, and rural areas have been left (relatively) behind. There used to be significant manufacturing hubs in Rochester (anchored by Kodak) and other mill towns, and upstate New York was an important agricultural producer. Rochester's GDP per worker was equal to New York City's from 1910 to 1970, but has grown more slowly since then and has converged with Savannah. With the relative decline of these upstate counties, New York City has increasingly diverged from the rest of the state, just as Atlanta has diverged from the rest Georgia.

These same patterns exist in many states, including Massachusetts where manufacturing cities in the western part of the state have diverged from the Boston metropolitan area. Figure 5 shows the GDP per worker of the top 20 MSAs by total GDP in 2018, all other MSAs, and all other counties.[8] For consistency, we define metropolitan areas based on recent areas but examining metro areas over long periods can be treacherous because metro areas today are not the same as the past. Las Vegas had a population of nearly zero in 1900, for example. Nonetheless, among these groups there was substantial convergence until approximately 1960. Since 1960, when the comparison is more reasonable, the top 20 MSAs have grown more quickly than other MSAs, while other MSAs and non-MSAs have grown at nearly the same pace. The top cities pulled away from the rest of the country.

Inequality within these groups shows a similar pattern to overall inequality. An extensive literature examines the convergence or divergence of metropolitan statistical areas in recent decades

[8]We define these MSAs using the BEA 2016 MSA definition, including all counties that are part of the MSA. The top 20 MSAs are the largest nominal GDP MSAs according to our GDP measure in 2018. They produce slightly less than 50 percent of all output in 2018. They include, in ascending order of GDP: Charlotte, Miami, Baltimore, Denver, San Diego, Detroit, Phoenix, Minneapolis, San Jose, Washington D.C., Seattle, Atlanta, Philadelphia, Boston, Houston, Dallas, San Francisco-Oakland, Chicago, Los Angeles, and New York.

(Moretti, 2012). Our data go much further back that this literature which focuses on wages rather than GDP. Keeping in mind that comparing current MSAs before 1950 is problematic, Figure 5 shows inequality among top MSAs), all other MSA, and non-MSA counties. Inequality declined within all groups from 1870 to about 1970. The overall inequality increase since then is partly explained by the divergence of the top MSA from other counties (shown in the left panel), increased inequality among top MSAs, and a smaller inequality increase among other MSAs and non-MSAs. Our results using GDP per worker are thus consistent with increasing wage dispersion documented in the MSA literature, but we show that a similar dispersion also occurred among non-MSAs. Divergence is not a phenomenon confined to cities.

3.3 Relationship between state GDP per worker and inequality

A different way to understand how across-state convergence and within-state divergence shaped the geography of production is to look at which states are the most unequal. Figure 6 shows how state GDP per worker relates to inequality within states. Before 1970, high GDP per worker states were the most equal. The relationship became mildly positive in the 1970s and strongly positive more recently so that the highest GDP per worker states are also the most unequal. Appendix Figure A-4 shows this relationship year by year; the slope is negative before 1958 and becomes positive in the 1970s. Many large urban centers were established largely for their role in trade which required a highly productive catchment area. New York City had high GDP because upstate New York had high GDP. After 1950, this relationship no longer holds as urban areas become decoupled from their hinterlands. We summarize this relationship as:

> **Result 4**: Before 1970 there is a negative relationship between state GDP per worker and inequality; more productive states were more unequal. In the 1970s the relationship becomes positive; more productive states were more unequal, particularly in more recent years.

4 Structural transformation in the U.S. and its contribution to inequality

We have documented how the county-level GDP per worker distribution evolved from 1870 to 2018 and described its main features. This long period encompassed a series of structural transformations as well. Our data, with their sectoral dimension, allows us to document this transformation and provide new evidence on how it has affected inequality. We start by summarizing the overall industrial composition. We then show how GDP per worker and industrial composition have evolved in Section 4.1. In Section 4.2, we formally decompose inequality by sector. There are

roughly four major industrial transformations since 1870 (see Appendix Figure A-12 for a comparison): (1) Agriculture's continuous decline from around 20 percent of GDP to nearly zero; (2) Manufacturing's increasing importance until 1960 and subsequent decline; (3) The slow increase in education and health services over the entire period; and (4) The rise of tradable services, including finance and business services such as software, since 1970. We show these industrial changes interact with geographic inequality in complex ways.

4.1 Industry shares

Figure 7 shows how the industry share of total county GDP varies from low GDP to high GDP per worker counties. The same pattern holds comparing the employment share across counties in Appendix Figure A-5. To reduce the complexity and for ease of presentation, we group industries which have a similar role in production or consumption into: Tradable services, which includes financial, insurance, and business and professional services; Non-tradable services, which includes recreation, personal, and retail services; Trade services including wholesale trade and transportation; and government, education, and health services.

Low GDP per worker counties used to be primarily agriculture counties. This relationship became less important over time until and by 2018 there is no strong relationship. Indeed, as a share of nominal GDP, the share agriculture now increases slightly with GDP per worker.

From 1870 to 1930, manufacturing became increasingly the activity of high GDP per worker counties. From 1980, manufacturing as a share was higher in middle GDP per worker counties than either low or high GDP per worker counties. High manufacturing concentration is no longer the sign of being a high GDP per worker county.

Instead, tradable services have become increasingly the main activity among high GDP per worker counties. These services have become increasingly concentrated in a few highly successful urban areas. Non-tradable services, on the other hand, used to be an approximately equal share of both low and high GDP counties. They have become more common in low-GDP counties where relatively little other productive activity occurs.

Government, education, and health services used to be an approximately equal share of high and low GDP per worker counties but have come to be a large share of the lowest GDP counties. We show in the next section they reduce inequality. Trade services (transportation and wholesale trade) were much more common in high GDP per worker counties in 1870 because many urban areas were defined by their role in trade. This concentration declined until, in 2018, trade services were fairly evenly spread.

We summarize these results as:

Result 5: Agriculture used to be the primary activity of the poorest counties. Now,

the poorer the county, the larger the share in government, education, and health.

Result 6: The path to riches has changed from manufacturing to tradable services. From 1870 to 1950 manufacturing became increasingly concentrated in the richest counties. The manufacturing share is now the highest in middle income counties, while the richest counties increasingly produce tradable services.

4.2 Industry contributions to inequality

We next turn to understanding how changes in industrial composition explain the overall changes in inequality across countries. To do this, we develop a sectoral decomposition of the Theil index. Shorrocks (1982) provides an axiomatic approach to breaking down income inequality into the components coming from different income sources. Because we consider GDP per worker, not total GDP, and we are interested in the contribution of various sectors to aggregate productivity inequality, the Shorrocks (1982) approach is not directly applicable. However, we develop a variant of it that provides a measure of each industry's contribution to inequality.

Denote the GDP share of each industry $s_c^k = Y_c^k/Y_c$, where $\sum_k s_c^k = 1$, summing over k constituent industries. Then the Theil index can be broken into k "pseudo-Theil" components giving the contribution of industry k to inequality:

$$T = \frac{1}{n}\sum_c \left((\sum_k s_c^k)\frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}} \right) = \sum_k \left(\frac{1}{n}\sum_c s_c^k \frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}} \right) = \sum_k \tilde{T}^k.$$

The industry contribution can alternatively be written also as:

$$\tilde{T}^k = \frac{1}{n}\sum_c s_c^k \frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}} = \frac{1}{n}\sum_c \hat{s}_c^k \frac{y_c^k}{\bar{y}} \ln \frac{y_c}{\bar{y}}$$

where $\hat{s}_c^k = N_c^k/N_c$ is the share of workers in sector *k* in county *c* (because $s_c^k = \hat{s}_c^k \frac{y_c^k}{y_c}$). Unlike a standard inequality index, it is possible for the pseudo-Theils to be less than zero.

We can rewrite the industry pseudo-Theils in a more intuitive way. There are two components that determine when the contribution of industry k to inequality will be positive or negative. One component is the mean share because small industries necessarily contribute relatively little on average to inequality. The other component is the covariance of the share with the Theil component

$\frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}}$. Formally, the psuedo-Theil can also be rewritten as:[9]

$$\tilde{T}^k = Cov[s_c^k, \frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}}] + \bar{s}^k T \tag{5}$$

where $\bar{s}^k = \sum_c s_c^k$ is the mean share of GDP in industry k.

Equation 5 helps provide intuition for the psuedo-Theils. If all counties have the same share, so the covariance is 0, then the contribution of industry k is simply its average share of the Theil. Mechanically, larger industries contribute more to inequality in the second term.

Industries with a larger positive covariance between the share and the Theil component also have a larger contribution to inequality in Equation 5's first term. The relationship between y_c and $\bar{s}^k$ is not entirely intuitive because the Theil component $\frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}}$ is not a monotonic transformation of y_c. However, if $y_c/\bar{y} > e^{-1} \approx 0.37$, then the Theil component is increasing in y_c. The implication is that, for all but the poorest counties, a positive relationship between y_c and s^k indicates a positive covariance between the share and the Theil component. Figure 7 shows the relationship between y_c and s^k directly.

Figure 8 shows the share each industry contributes to GDP per worker inequality ($\tilde{T}^k/T$). Appendix Figure A-6 shows the corresponding direct industry contribution $\tilde{T}^k$. We show all industries except real estate. Several important points stand out.

First, the rapid decline in inequality from 1870 to 1970 is explained by a broad inequality decline across several industries including manufacturing, transportation, communication, utilities, wholesale trade, and government. But the sharp decline from 1930 to 1970 is almost entirely explained by the diminishing contribution of manufacturing. While manufacturing's contribution to inequality increases from 1870 to 1930, that is because manufacturing inequality is relatively constant while other industries declined (see Appendix Figure A-6). Manufacturing went from accounting for approximately 40 percent of inequality before 1940 to a much smaller contribution after 1970, although its contribution rises somewhat around 2000. After 1970, manufacturing contributed to the increase inequality as well, but not on the scale of its previous contribution.

Wholesale trade and transportation's contribution to inequality also declined significantly from 1870 and helped explain the fall in inequality to 1970, although their relatively small share reduces the impact. Tradable services' contribution to inequality increases noticeably in recent decades.

[9]Define $x_c = \frac{y_c}{\bar{y}} \ln \frac{y_c}{\bar{y}}$ and note that $T = \bar{x} = \frac{1}{n} \sum x_c$. Then

$$\begin{aligned} Cov[s_c^k, x_c] &= \frac{1}{n} \sum (s_c^k - \bar{s}^k)(x_c - \bar{x}) \\ &= \frac{1}{n} \sum (s_c^k x_c) - (s_c^k \bar{x}_c) - (\bar{s}^k x_c) + (\bar{s}^k \bar{x}_c) \\ &= \tilde{T}^k - \bar{s}^k T - \bar{s}^k T + \bar{s}^k T = \tilde{T}^k - \bar{s}^k T. \end{aligned}$$

Second, agriculture and mining inequality have spikes in their contribution to nominal inequality largely explained by price swings. These spikes come from rapid changes in prices which rapidly change the nominal GDP of agricultural or resource extraction counties. Because counties where agriculture or mining are important also typically have relatively small populations, commodity price swings can have a big impact on their GDP per worker and thus inequality. The spike in agriculture's inequality contribution in the early 1970s comes from the rapid increase in grain and oil seed prices from 1971 to 1974. Oil price fluctuations in the 1970s and 1980s, as well as the more recent fracking boom, explain the changing contribution of mining to inequality.

Fourth, competing trends explain the rise in inequality since 1970. The tradable services' contribution to inequality has been increasing. Together, they now explain more than 40 percent of all GDP inequality. Manufacturing's contribution to inequality increased since 1970. Meanwhile, government services which had increased inequality before 1950, instead decrease inequality after 1970.

The industry concentration and productivity dispersion help explain industry inequality contributions. One way to measure concentration across counties is the isolation index shown in Figure 9. Isolation measures the average industry share of GDP where the average dollar of production in that industry takes place.[10] High values indicate that most production in that industry occurs in counties where that industry represents a large fraction of local production, so the industry is highly concentrated; low values indicate production occurs relatively evenly across counties.

Three trends are evident that help explain the change in production: (1) Agriculture used to be primarily produced in counties that mostly produced agriculture. Today the average dollar of agriculture is produced in counties where it is around 10 percent of county GDP. It is worth remembering that, even in most agriculture-intensive counties, agriculture is no longer the largest industry. (2) Manufacturing concentration first increases and then decreases. At the peak, manufacturing was highly concentrated in counties where it represented 30 percent or more of production. The subsequent decline in concentration partly explains the decline in manufacturing's contribution to inequality. (3) Tradable services used to be evenly spread out and had a small share overall. They have become increasingly concentrated, so the average dollar of tradable services is now produced in a county where they represent 25 percent of GDP.[11]

Figure 10 shows how GDP per worker inequality has changed for major industries, using the

[10]The isolation index, with N denoting national production, is:

$$\text{Isolation}_t^k = \sum_c \left(\frac{Y_{ct}^k}{Y_{Nt}^k}\right)\left(\frac{Y_{ct}^k}{Y_{ct}}\right)$$

[11]These trends are similar to Desmet and Fafchamps (2006) who examine employment concentration across counties. They find an increasing concentration of jobs in services between 1970 and 2000, while manufacturing and farming employment become more evenly spread.

the Theil index. Because industries with very low production may have unrealistically high or low productivity, we exclude the top and bottom 0.5 percent of industry GDP per worker counties.

Three facts stand out: First, manufacturing productivity inequality increased until 1940, decreased from 1940 to 1970, and slightly increased since then. The combined decrease in manufacturing concentration and the decrease in the manufacturing productivity dispersion account for the reduction in manufacturing's contribution to inequality from 1930 to 1970. Second, agriculture productivity inequality increased up to 1940, decreased from 1940 until 1970, and then substantially increased since 1970. However, because agriculture is a small share of GDP, the increasing agriculture productivity inequality has little impact on overall inequality. Third, there has been a small increase in tradable services' productivity inequality since 1990.

While there are many overlapping changes, we summarize the most important ones as:

> **Result 7:** Manufacturing's contribution to inequality is the largest before 1960 and its decline is the main explanation for the fall in inequality from 1930 to 1970. Manufacturing's declining contribution to inequality is explained by both greater geographical diffusion and falling cross-county productivity inequality. Finance and professional services contribute to the rise in GDP per worker inequality, mostly because they have become more concentrated.

Finally, Figure 7 and Result 5 show that government services have become a larger share of the poorest counties' economies since 1930. Using the psuedo-Theil, Figure 8 shows that government services directly reduce inequality. We summarize this as:

> **Result 8:** After 1950, government services reduce inequality. Before 1950, government services increased inequality.

5 Convergence, divergence, and growth

We have documented a large convergence across U.S. counties from 1870 to 1970 and a subsequent divergence. In this section, we examine important these changes were for overall growth and some partial answers for what changed.

We first introduce a simple Productivity Gap Index (PGI) which measures how important productivity differences are for GDP. [12]

$$\text{PGI} = \frac{1}{GDP} \sum_{i=1}^{N} L_i(z - y_i)1(y_i\text{<z}) \tag{6}$$

[12]The PGI is inspired by, and very similar to, the "Poverty Gap Index" used in poverty and development research to account for inequality among the poor and defined as $P_1 = \frac{1}{N}\sum_{i=1}^{N} \frac{(z-y_i)1(y_i<z)}{z}$ where z is the poverty line, N is the population, and y_i is the consumption or income of people below the poverty line.

where $(z - y_i)$ is the productivity gap between county y_i and an appropriately defined leader z, multiplied by the county employment L_i to obtain the GDP cost of gap. The PGI thus measures how much GDP would increase if all $y_i < z$ were brought up to z and multiplying GDP by $(1 + PGI)$ gives what the hypothetical frictionless GDP would be. It thus measures the combination of housing (Hsieh and Moretti, 2019), transportation (Donaldson and Hornbeck, 2016), and other geographic missallocations or frictions (Hsieh and Klenow, 2009, 2010) that reduce productivity.

Figure 11 shows the PGI with the leader z defined at the 99th percentile, as well as several others for comparison.[13] Unsurprisingly, bringing low productivity counties up to leaders increase GDP; what is important is how much inequality matters and that it changes over time. In fact, some regional inequality is unimportant for total productivity. Bringing below median counties up to the median has a negligible effect on total GDP. Low productivity counties typically have small populations and their economies are now dominated by education, health, and government services (Figure 7) which have a lower productivity dispersion (Figure 10). On the other hand, bringing counties above the median, particularly the larger ones, up to the productivity frontier has a large effect on GDP. In 1870, real GDP would have hypothetically been 79 percent larger with no productivity gap. The gap fell to only 21 percent in 1970. As inequality has increased since 1970, so too has the productivity gap, reaching 42 percent in 2018.

Altogether, a substantial portion of total growth is explained by changing productivity inequality. Real GDP per worker increased at an annual rate of 1.7 percent from 1870 to 1970. If the productivity gap had stayed the same, real GDP per worker would have increased by only 1.3 percent annually.[14] From 1900 to 1970, actual real GDP per worker growth was 1.73 percent annually and would have been 1.34 percent without convergence, suggesting that this calculation is robust to the starting point. In total, real GDP per worker would have been 176 percent smaller in 1970 which is 39 percent of the total GDP per worker growth that actually occurred between 1870 and 1970. After 1970, real annual GDP per worker growth was 1.52 percent annually from 1970 to 2018. Growth would have been 22 percent higher, 1.86 percent annually, if the productivity gap had not increased and real GDP per worker would have been 35 percent larger in 2018. Using a lower frontier reduces the gains, of course, but not much. Using the 95th percentile, annual real GDP per worker growth would have been 1.46 annually without convergence and GDP per worker 120 percent smaller in 1970. While much has been made of the productivity growth slowdown, particularly starting after 2000 (Syverson, 2017), our results show slowing aggregate productivity

[13] In calculating the percentiles, we exclude counties from being in the top percentile that have a mining share greater than 30 percent and have fewer than 10,000 people. Removing smaller counties and mining counties makes the PGI less sensitive to measurement error and price spikes and so closer to a measurement of the productivity frontier. The PGI not making this exclusion is somewhat larger in all years, but the overall trends are similar and the contribution to growth is similar.

[14] We compare GDP holding the PGI constant. Defining Y_t as real GDP at t, and PGI_t and $PGI_{t'}$ as the Productivity Gap at t and t', then real GDP if the gap at t were actually $PGI_{t'}$ is: $Y_t \frac{(1+PGI_t)}{(1+PGI_{t'})}$.

is mechanically explained by changing convergence.

What explains why convergence has reversed? We first show a large decline in people moving to highly productive areas. Standard convergence theory suggests that factor mobility should drive convergence as people and capital move to where they will have their highest marginal return. Figure 12 shows the relationship between GDP per worker and population growth over the next decade. During the period of rapid GDP per worker convergence up to 1970, the higher GDP per worker counties' population grew much more quickly, partly because they were a destination for international immigrants. However, starting in the 1960s, the population of high GDP per worker counties grew only slightly faster than other counties, and grew *slower* during the 1980s and early 1990s as people moved out of the highest GDP per worker places (Appendix Figure A-7 shows year by year linear regression coefficients). While the relationship turned modestly positive again starting in the mid-1990s, population movement no longer strongly favored convergence.

Recent research (Glaeser and Gottlieb, 2009; Hsieh and Moretti, 2019; Ganong and Shoag, 2017) has emphasized the importance of housing costs in preventing people from moving to high productivity areas. An individual worker, even if her wage and marginal product is higher in a more productive area, may not find moving to that area improves her welfare because of the high cost of housing. Similarly, the presence of frictions and limited supply of factors such as land may speed or slow population movement's contribution to convergence (Desmet and Rappaport, 2017). While our results do not provide a reason for the shift, they show that pro-convergence population growth was large during the period of convergence and has since diminished or has favored divergence.

Human capital also converged through the 1970s but its role has become more complicated in recent decades (Berry and Glaeser, 2005). Appendix Figure A-8 shows county inequality in education using three measures: the share literate before 1930, years of education from 1940 to 2010, and the share with a college degree (or more) from 1970 to 2020. Appendix Figure A-9 shows the same measures across the top MSAs, other MSAs, and all other counties. There was rapid education convergence from 1870 until 1970 which may explain the rapid GDP per worker convergence until 1970. Indeed, the education convergence may have been a precondition to manufacturing's geographical diffusion. By 1990, average years of education inequality had diminished substantially and was approaching zero. Since 1970, inequality in the share of the population with a college degree has diminished slightly and remains mostly within-state inequality.

While Figure A-8 suggests education inequality continued to decline after 1970, the top MSAs have been pulling away from other MSAs and non-MSAs in college share. In 1970, the college share in top MSAs was 2 percentage points higher than in other MSAS; by 2020 it was 9 percentage points higher. Thus, while Figure A-8 shows an overall convergence across all counties in education when treating counties as the unit, Figure A-9 shows that the top MSAs have increasingly pulled in the most educated workers as they have come to dominate tradable services. Tradable

services' increasing geographic concentration may have been aided by better communication technology that favored regional specialization in high-skill services and so contributed to a growing and more unequal college wage premium (Eckert, 2019).

Other research points to the increasing importance of agglomeration economies (Glaeser and Gottlieb, 2009; Moretti, 2012; Diamond, 2016). Figure 13 shows how productive MSAs other than the top MSA and non-MSAs are compared to the top MSAs. Top MSAs and other MSAs had nearly the same manufacturing labor productivity from around 1900 to 1980 but have since diverged suggesting that agglomeration economies in manufacturing are growing. Even more pronounced, the top MSAs have become much more productive in producing tradable services than other areas. Tradable services are also becoming more concentrated in high GDP per worker areas (Result 6). These results, taken together, suggest that the interaction between agglomeration effects and industrial shifts is contributing to the recent divergence.

We summarize these results as:

> **Result 9:** Population growth and education used to be strongly pro-convergence, but after 1970 became neutral or anti-convergence. At the same time, agglomeration effects in manufacturing and tradable services appear to have increased.

6 Conclusion

We use our novel data to document a fall in GDP per worker inequality from 1870 to about 1970 and then a subsequent rise. This rise is mostly due to increasing within state inequality as leading metropolitan areas have converged while diverging both from other cities and rural areas. Inequality is no longer primarily regional. Indeed, the highest GDP per worker states are now the most unequal. Atlanta looks more like New York City, but upstate New York looks more like the rest of Georgia.

Our data let us analyze how sectoral transformations contribute to this convergence and then divergence. Manufacturing's geographic diffusion underlies much of the decline in productivity inequality until the 1970s, while tradable services' growth and increasing concentration contribute to the rise of GDP per worker inequality after that. Meanwhile, there are equalizing forces as government, education, and health have come to be a larger share of the least productive counties and government services reduce inequality directly.

Are we still growing together? No, but why and how are complicated. The conditions that led to the rapid convergence until 1970 no longer apply. Population mobility slowed, and it is no longer directed towards the highest GDP per worker counties. Education converged across all counties, but the top MSAs have acquired a large advantage in the percentage of college educated workers. Manufacturing's geographic diffusion slowed and then partially reversed. And new agglomeration

economies appear to be driving greater concentration in tradable services. Our results add novel evidence to the rich literature on regional and subregional development and they emphasize the importance of further research on the impediments to growing together.

References

Annual Reports of the Postmaster General. 1900. *Report of the Postmaster General*. Washington, D.C.: Government Printing Office.

Atack, Jeremy, Robert A. Margo, and Paul W. Rhode. 2022. "Industrialization and urbanization in nineteenth century America." *Regional Science and Urban Economics* 94:103678. URL https://www.sciencedirect.com/science/article/pii/S0166046221000387. Urban Economics and History.

Austin, Benjamin, Edward Glaeser, and Lawrence H. Summers. 2018. "Saving the heartland: Place-based policies in 21st Century Americas." *Brookings Papers on Economic Activity* (Spring):151–255.

Autor, David, David Dorn, and Gordon Hanson. 2021. "On the Persistence of the China Shock." *Brookings Papers on Economic Activity* (Fall):381–447. URL https://www.brookings.edu/articles/on-the-persistence-of-the-china-shock/.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi. 2020. "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure." *American Economic Review* 110 (10):3139–83. URL https://www.aeaweb.org/articles?id=10.1257/aer.20170011.

Autor, David H., David Dorn, and Gordon H. Hanson. 2016. "The China Shock: Learning from Labor Market Adjustment to Large Changes in Trade." *Annual Review of Economics* 8 (1):205–240. URL https://www.nber.org/papers/w21906.

Aysheshim, Kirubel D., Jacob R. Hinson, , and Sharon D. Panek. 2020. "A Primer on Local Area Gross Domestic Product Methodology: First Official Release of Gross Domestic Product by County, 2001–2018." *Survey of Current Business* 100 (3). Available: https://apps.bea.gov/scb/issues/2020/03-march/0320-county-level-gdp.htm.

Barro, Robert J. 2015. "Convergence and Modernisation." *The Economic Journal* 125 (585):911–942. URL https://doi.org/10.1111/ecoj.12247.

Barro, Robert J. and Xavier Sala-i-Martin. 1991. "Convergence across States and Regions." *Brookings Papers on Economic Activity* 22 (1):107–182.

———. 1992. "Convergence." *Journal of Political Economy* 100 (2):223–251. URL http://www.jstor.org/stable/2138606.

———. 2004. *Economic Growth*. Cambridge: MIT Press, 2nd ed. URL http://www.jstor.org/stable/2138606.

Bernard, Andrew B. and Charles I. Jones. 1996. "Productivity and convergence across U.S. States and industries." *Empirical Economics* 21 (1):113–135. URL https://doi.org/10.1007/BF01205496.

Berry, Christopher R. and Edward L. Glaeser. 2005. "The divergence of human capital levels across cities." *Papers in Regional Science* 84 (3):407–444. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1435-5957.2005.00047.x.

Blanchard, Olivier Jean and Lawrence F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 23 (1):1–76. URL http://ideas.repec.org/a/bin/bpeajo/v23y1992i1992-1p1-76.html.

Breinlich, Holger, Gianmarco I.P. Ottaviano, and Jonathan R.W. Temple. 2014. "Chapter 4 - Regional Growth and Regional Decline." In *Handbook of Economic Growth*, *Handbook of Economic Growth*, vol. 2, edited by Philippe Aghion and Steven N. Durlauf. Elsevier, 683 – 779. URL http://www.sciencedirect.com/science/article/pii/B9780444535405000045.

Bureau of Municipal Research of Philadelphia. 1916. *Comparative Salary Data*.

Bureau of the Census. 1950. "Alphabetical Index of Occupations and Industries: 1950." Available: https://usa.ipums.org/usa/resources/volii/Occupations1950.pdf, accessed 7 April 2017.

Carson, Daniel. 1949. *Studies in Income and Wealth*, chap. Changes in the Industrial Composition of Manpower since the Civil War. NBER, 46–150.

Carter, Susan B. 2006. *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*, chap. National income, by industry group: 1929–2002. Cambridge University Press, Table Ca35–53. URL http://dx.doi.org/10.1017/ISBN-9780511132971.Ca1-26.

Caselli, Francesco and Wilbur John Coleman, II. 2001. "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation." *Journal of Political Economy* 109 (3):pp. 584–616. URL http://www.jstor.org/stable/10.1086/321015.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* URL http://qje.oxfordjournals.org/content/early/2014/10/16/qje.qju022.abstract.

Comin, Diego, Danial Lashkari, and Martí Mestieri. 2021. "Structural Change with Long-Run Income and Price Effects." *Econometrica* 89 (1):311–374. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA16317.

Cowell, F.A. 2000. "Chapter 2 Measurement of inequality." Elsevier, 87–166. URL https://www.sciencedirect.com/science/article/pii/S1574005600800056.

Desmet, Klaus and Marcel Fafchamps. 2006. "Employment concentration across U.S. counties." *Regional Science and Urban Economics* 36 (4):482–509. URL https://www.sciencedirect.com/science/article/pii/S0166046206000330.

Desmet, Klaus and Jordan Rappaport. 2017. "The settlement of the United States, 1800–2000: The long transition towards Gibrat's law." *Journal of Urban Economics* 98:50–68. URL https://www.sciencedirect.com/science/article/pii/S0094119015000224. Urbanization in Developing Countries: Past and Present.

Desmet, Klaus and Esteban Rossi-Hansberg. 2009. "Spatial growth and industry age." *Journal of Economic Theory* 144:2477–2502.

———. 2014. "Spatial Development." *American Economic Review* 104 (4):1211–43. URL http://www.aeaweb.org/articles?id=10.1257/aer.104.4.1211.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106 (3):479–524. URL https://www.aeaweb.org/articles?id=10.1257/aer.20131706.

Donaldson, Dave and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *The Quarterly Journal of Economics* 131 (2):799–858. URL https://doi.org/10.1093/qje/qjw002.

Easterlin, Richard. 1960. *Trends in the American Economy in the Nineteenth Century*, vol. 24. Princeton University Press. URL http://www.nber.org/chapters/c2475.

Eckert, Fabian. 2019. "Growing Apart: Tradable Services and the Fragmentation of the U.S. Economy."

Eckert, Fabian and Michael Peters. 2022. "Spatial Structural Change." Working Paper 30489, National Bureau of Economic Research. URL http://www.nber.org/papers/w30489.

Eichengreen, Barry and Poonam Gupta. 2013. "The two waves of service-sector growth." *Oxford Economic Papers* 65:96–123.

Fulford, Scott, Ivan Petkov, and Fabio Schiantarelli. 2016. "Does It Matter Where You Came From? Ancestry Composition and Economic Performance of US Counties, 1850 - 2010." URL http://ssrn.com/abstract=2612312.

———. 2020. "Does It Matter Where You Came From? Ancestry Composition and Economic Performance of US Counties, 1850 - 2010." *Journal of Economic Growth* 25 (3). URL https://doi.org/10.1007/s10887-020-09180-9.

Fulford, Scott L. 2015. "How important are banks for development? National banks in the United States 1870–1900." *Review of Economics and Statistics* 97 (5):921–938. URL http://fmwww.bc.edu/ec-p/wp753.pdf.

Gallman, Robert E. and Thomas J. Weiss. 1969. *The Service Industries in the Nineteenth Century*. National Bureau of Economic Research and UMI, 287–381. URL http://www.nber.org/chapters/c1205.

Ganong, Peter and Daniel Shoag. 2017. "Why Has Regional Convergence in the U.S. Declined?" *Journal of Urban Economics* 102:76–90. URL https://doi.org/10.1016/j.jue.2017.07.002.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. 2013. "Human Capital and Regional Development." *The Quarterly Journal of Economics* 128 (1):105–164. URL http://EconPapers.repec.org/RePEc:oup:qjecon:v:128:y:2013:i:1:p:105-164.

Giannone, Elisa. 2019. "Skill-Biased Technical Change and Regional Convergence." URL https://sites.google.com/view/elisagiannone/research/working-papers.

Glaeser, Edward L. and Joshua D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4):983–1028. URL http://www.aeaweb.org/articles?id=10.1257/jel.47.4.983.

Goldin, Claudia and Lawrence F. Katz. 2010. *The Race between Education and Technology*. Cambridge, MA: Harvard University Press.

Goldin, Claudia and Robert A. Margo. 1992. "The Great Compression: The Wage Structure in the United States at Mid- Century." *The Quarterly Journal of Economics* 107 (1):1–34. URL http://www.jstor.org/stable/2118322.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2):994–1028. URL http://www.aeaweb.org/articles?id=10.1257/aer.102.2.994.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi. 2014. "Growth and Structural Transformation." In *Handbook of Economic Growth*, *Handbook of Economic Growth*, vol. 2, edited by Philippe Aghion and Steven N. Durlauf. Elsevier, 855 – 941. URL http://www.sciencedirect.com/science/article/pii/B9780444535405000069.

Higgins, Matthew J., Daniel Levy, and Andrew T. Young. 2006. "Growth and Convergence across the United States: Evidence from County-Level Data." *The Review of Economics and Statistics* 88 (4):671–681.

Hsieh, Chang-Tai and Peter Klenow. 2010. "Development Accounting." *American Economic Journal: Macroeconomics* 2:207–223.

Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4):1403–1448. URL http://qje.oxfordjournals.org/content/124/4/1403.full.pdf.

Hsieh, Chang-Tai and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11 (2):1–39. URL https://www.aeaweb.org/articles?id=10.1257/mac.20170388.

James, John A. 1976. "Banking Market Structure, Risk, and the Pattern of Local Interest Rates in the United States, 1893-1911." *The Review of Economics and Statistics* 58 (4):453–462. URL http://www.jstor.org/stable/1935877.

Kleinman, Benny, Ernest Liu, and Stephen J. Redding. 2023. "Dynamic Spatial General Equilibrium." *Econometrica* 91 (2):385–424. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA20273.

Leven, Maurice. 1925. *Income in the Various States: Its Sources and Distribution, 1919, 1920, and 1921*. NBER. URL http://papers.nber.org/books/leve25-1.

Lindert, Peter H. and Jeffrey G. Williamson. 2016. *Unequal Gains: American Growth and Inequality since 1700*. Princeton, NJ: Princeton University Press.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding. 2012. "Urbanization and Structural Transformation." *The Quarterly Journal of Economics* 127 (2):535–586. URL https://doi.org/10.1093/qje/qjs003.

Minnesota Population Center. 2011. "National Historical Geographic Information System: Version 2.0." Minneapolis, MN, University of Minnesota. Available: http://www.nhgis.org, accessed 1 August 2014.

Moretti, Enrico. 2012. *The New Geography of Jobs*. Houghton Mifflin.

Nineteenth Annual Report of the Commissioner of Labor. 1905. *Wages and Hours of Labor 1904*. Government Printing Office.

Piketty, Thomas and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *The Quarterly Journal of Economics* 118 (1):1–41. URL https://doi.org/10.1162/00335530360535135.

Pisati, Maurizio. 2007. "SPMAP: Stata module to visualize spatial data." Statistical Software Components, Boston College Department of Economics. URL https://ideas.repec.org/c/boc/bocode/s456812.html.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2019. "Integrated Public Use Microdata Series: Version 9.0 [dataset]." URL https://doi.org/10.18128/D010.V9.0.

Schwartz, Charles F. and Robert E. Graham, Jr. 1956. *Personal income by States since 1929: A supplement to the Survey of Current Business*. Washington, D.C.: U.S. Government Printing Office.

Shorrocks, A. F. 1982. "Inequality Decomposition by Factor Components." *Econometrica* 50 (1):193–211. URL http://www.jstor.org/stable/1912537.

Sutch, Richard. 2006. *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*, chap. Gross domestic product: 1790-2002 [Continuous annual series]. Cambridge University Press, Table Ca9–19. URL http://dx.doi.org/10.1017/ISBN-9780511132971.Ca1-26.

Syverson, Chad. 2017. "Challenges to Mismeasurement Explanations for the US Productivity Slowdown." *Journal of Economic Perspectives* 31 (2):165–86. URL https://www.aeaweb.org/articles?id=10.1257/jep.31.2.165.

Theil, Henri. 1967. *Economics and information theory*. Amsterdam: North-Holland Publishing Company.

United States Bureau of Labor Statistics. 1919. *Wages and Hours in the Coal Mining Industry in 1919*. Goverment Printing Office.

United States Department of Commerce. 1993. *National Income and Product Accounts of the United States*, vol. Volume 1, 1929-58. Washington, D.C.: U.S. Government Printing Office. URL http://fraser.stlouisfed.org/docs/publications/nipab/nipa_1929_1958.pdf.

———. 2001. *National Income and Product Accounts of the United States, 1929-1997*, vol. Volume 2. Washington, D.C.: U.S. Government Printing Office. URL https://fraser.stlouisfed.org/docs/publications/nipab/nipa_1929_97_v2.pdf.

United States Department of Commerce Bureau of the Census. 2012. "County and City Data Book Consolidated File: County Data, 1947-1977 (ICPSR07736-v2)." Inter-university Consortium for Political and Social Research [distributor]. URL http://doi.org/10.3886/ICPSR07736.v2.

United States Department of Labor. 1934. *History of Wages in the US from Colonial Times to 1928*. Washington, D.C.: Goverment Printing Office, republished by gale research company, detroit, 1966. ed.

U.S. Bureau of Economic Analysis. 2016. "Local Area Personal Income Accounts Tables CA5 and CA5N." URL http://www.bea.gov/regional/downloadzip.cfm.

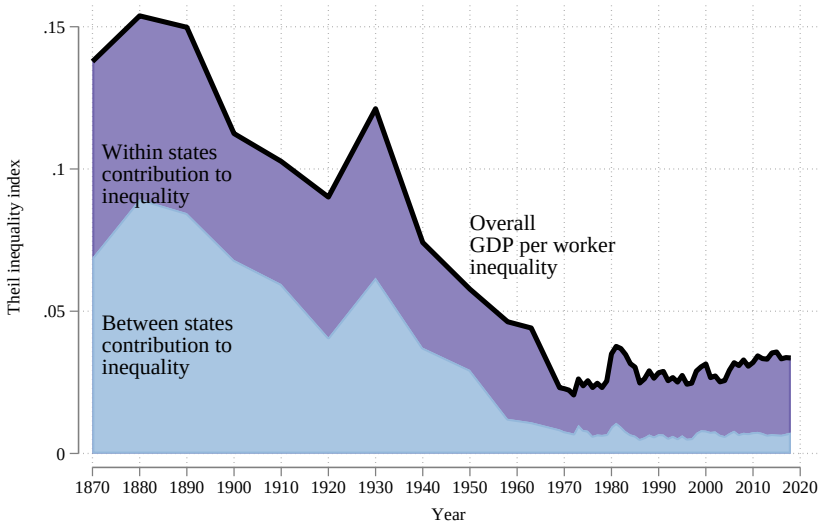
U.S. Census Bureau. 1975. *Historical Statistics of the United States, Colonial Times to 1970, Bicentenial Edition, Part 1*. Washington, D.C.: U.S. Government Printing Office. URL https://www.census.gov/prod/www/statistical_abstract.html.

Weber, Warren E. 2000. "Disaggregated Call Reports for U.S. National Banks, 1880-1910. Research Department, Federal Reserve Bank of Minneapolis." Research Department, Federal Reserve Bank of Minneapolis. Available: http://research.mpls.frb.fed.us/research/economists/wewproj.html, accessed 30 April 2013.

Wright, Gavin. 2006. *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*, chap. Mineral operations - establishments, receipts, value added, employment, and expenses: 1880-1997. Cambridge University Press, Table Db1–11. URL http://dx.doi.org/10.1017/ISBN-9780511132971.Db1-154.

Young, Andrew T., Matthew J. Higgins, and Daniel Levy. 2008. "Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data." *Journal of Money, Credit and Banking* 40 (5):1083–1093.

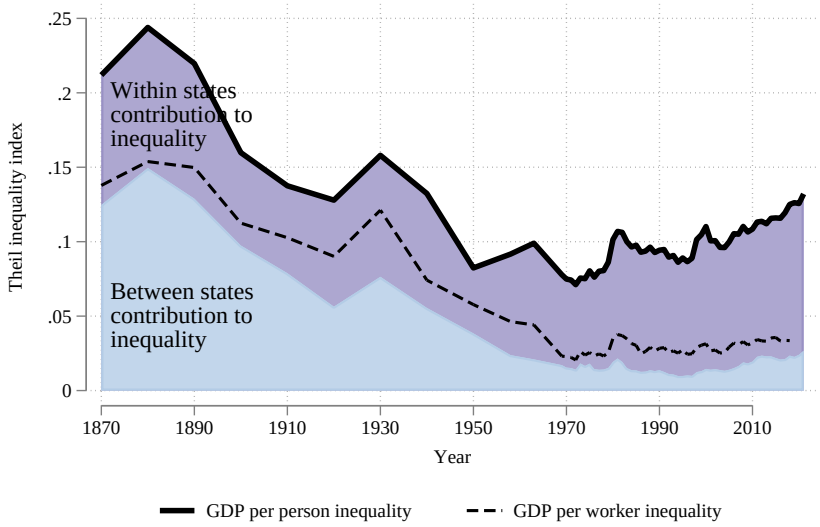
Figure 1: GDP per worker inequality



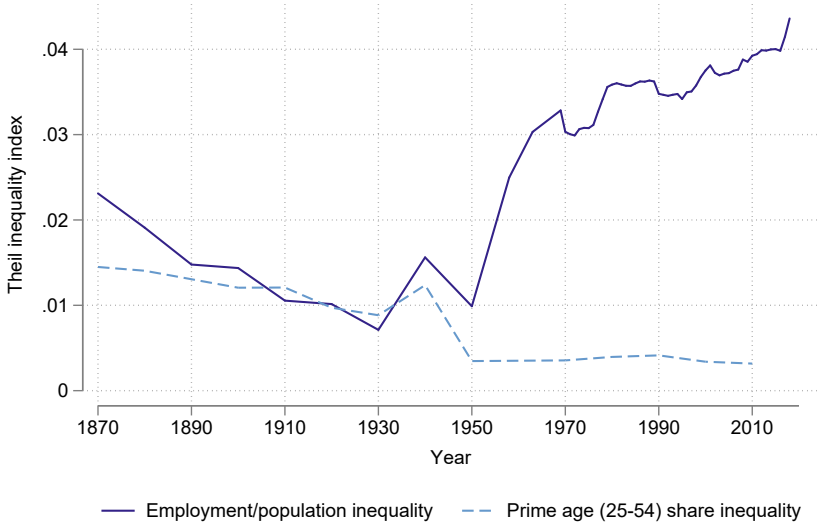
Notes: Main sample excludes counties with mining share of GDP greater than 40 percent, utilities share greater than 25 percent, or have a populations smaller than 2500 (see Section 2.3).

Figure 2: GDP per person inequality and dependency inequality

Panel (A): GDP per person inequality



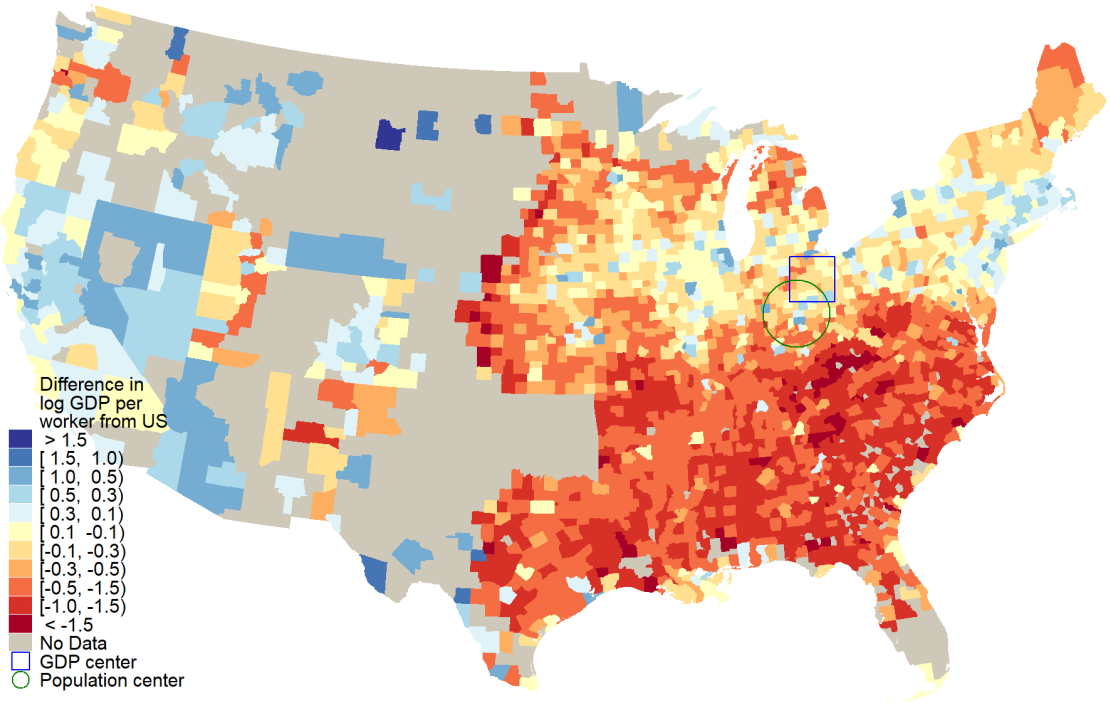
Panel (B): Employment and prime-age population share inequality



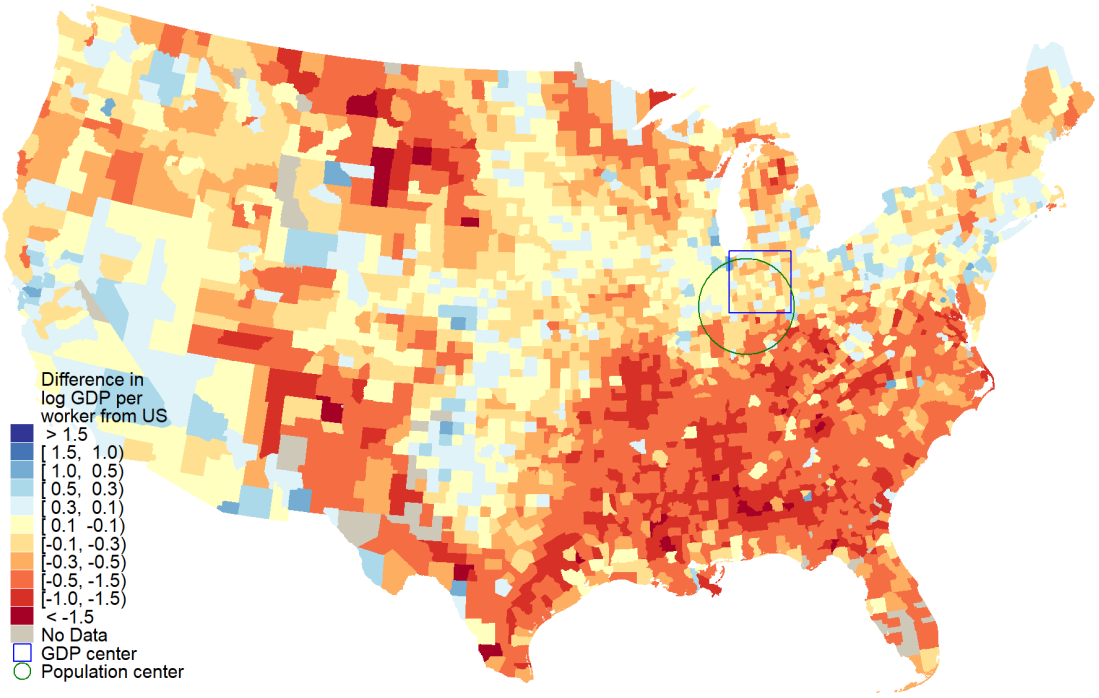
Notes: Common deflator. Employment in 1950 or before is based on Census occupation that includes market-based activities. Starting in 1969 it is the BLS jobs definition, so the change between 1950 and 1969 may be from a definitional change. Employment /population is for all ages. Prime age is the share age 25-54 of the population. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 3: The geographic distribution of GDP per worker: 1880 and 1920

1880



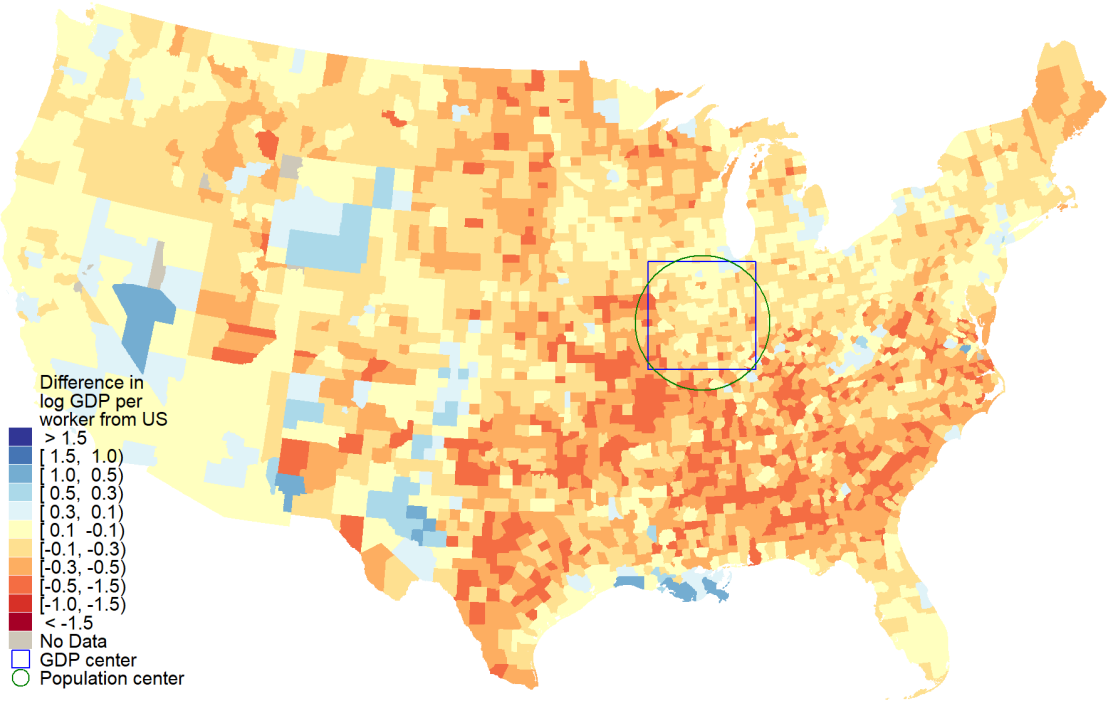
1920



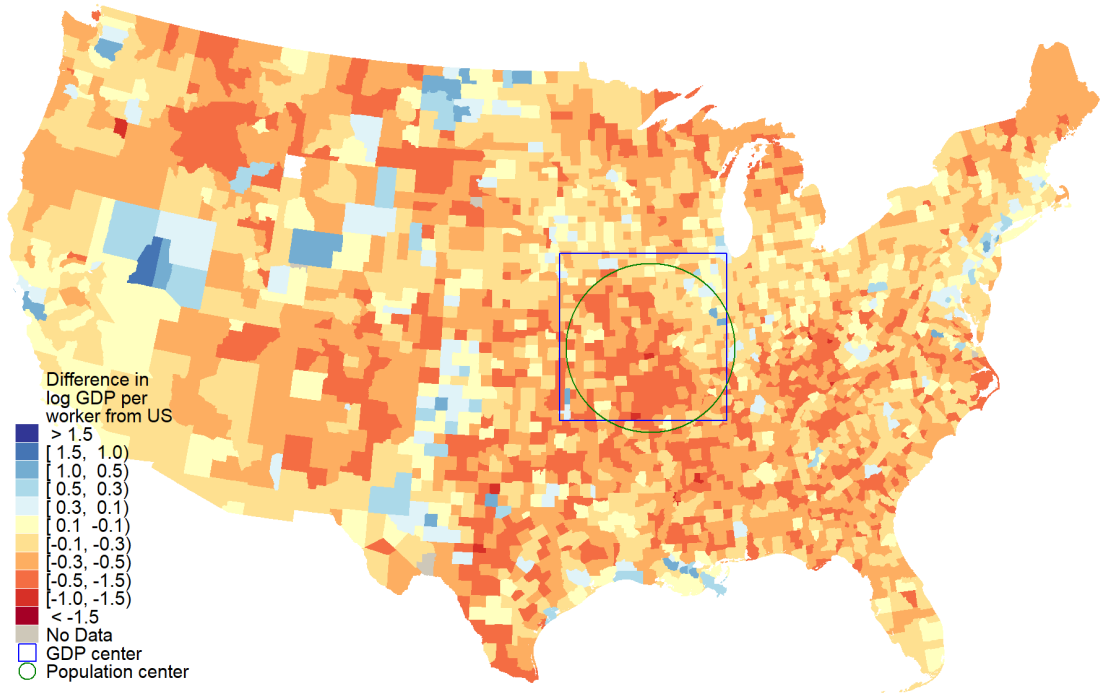
Notes: Shading in each county is by the difference in log county GDP per worker and log U.S. GDP per worker. The circle is the population center of the U.S. and is proportional to population in 2010. The square shows the GDP center of the U.S. and is proportional to 2010 real GDP per person. Uses county definitions from Minnesota Population Center (2011) and mapping software from Pisati (2007). The sample excludes counties with population less than 2500.

Figure 4: The geographic distribution of GDP per worker: 1970 and 2018

1970

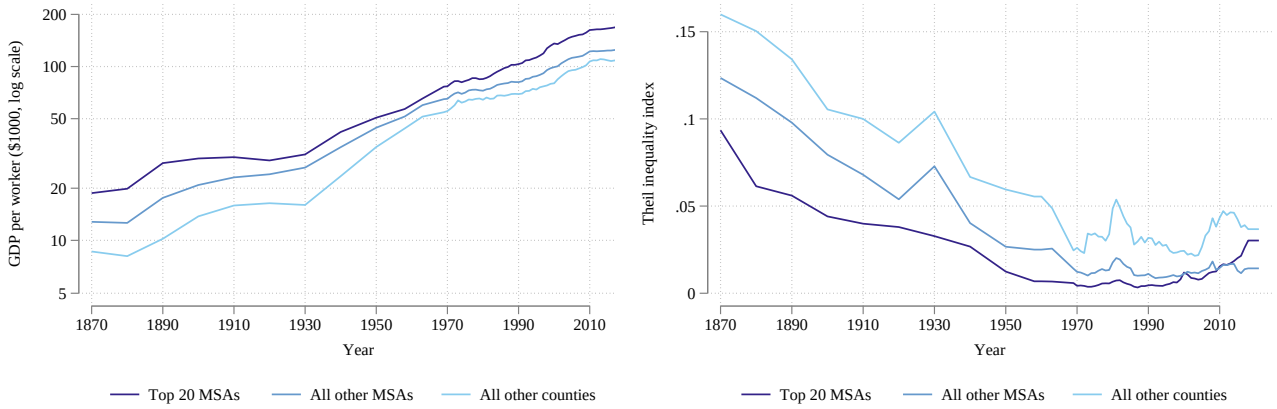


2018



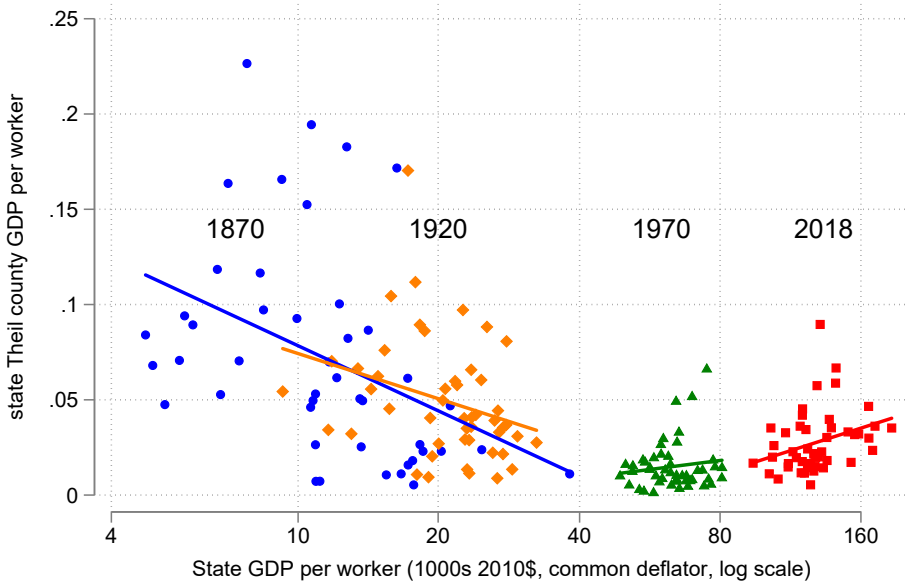
Notes: Shading in each county is by the difference in log county GDP per worker and log U.S. GDP per worker. The circle is the population center of the U.S. and is proportional to population in 2010. The square shows the GDP center of the U.S. and is proportional to 2010 real GDP per person. Uses county definitions from Minnesota Population Center (2011) and mapping software from Pisati (2007). The sample excludes counties with population less than 2500.

Figure 5: GDP per worker and inequality across metropolitan and non-metropolitan areas



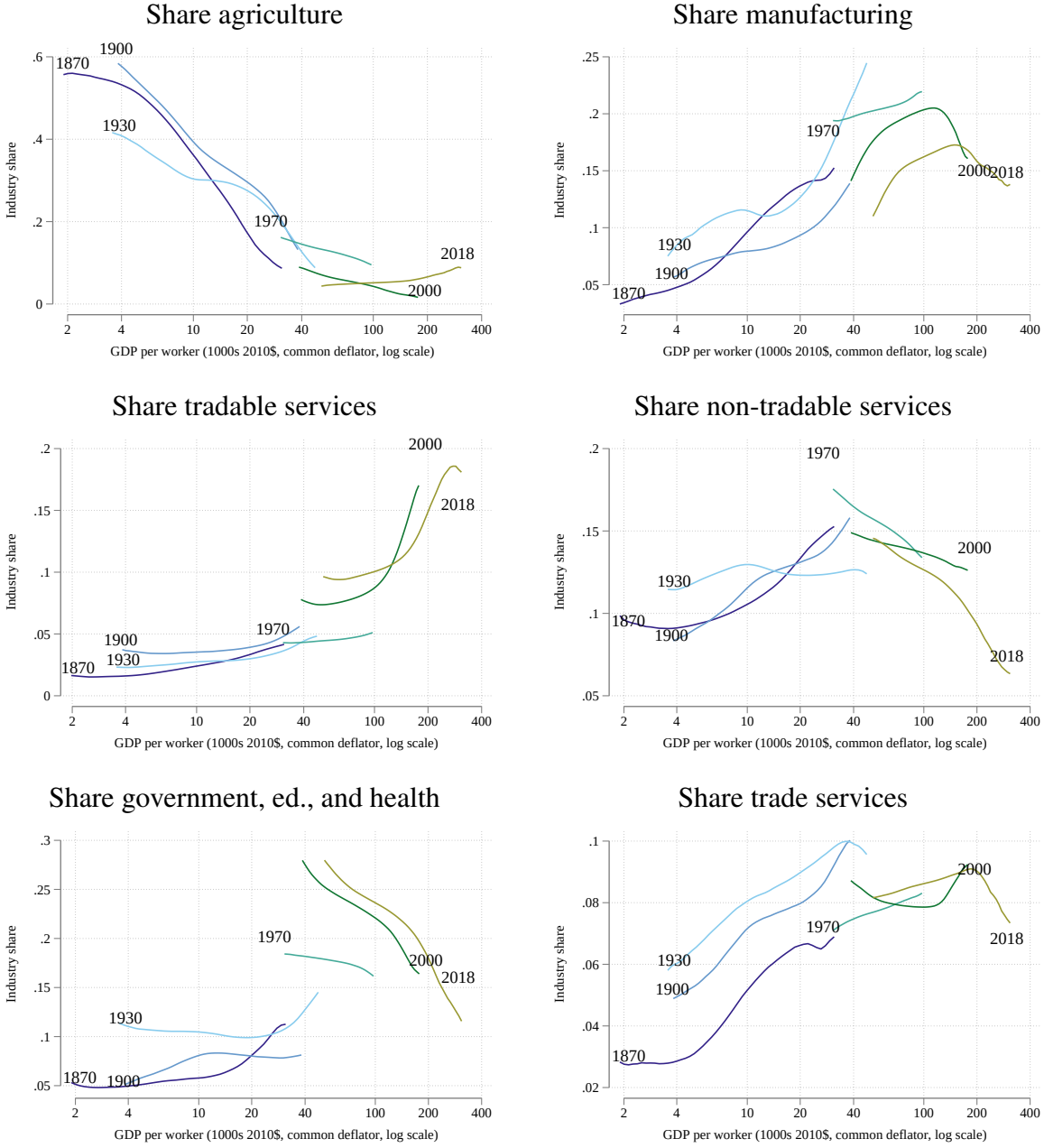
Notes: We use the counties in MSAs as defined by BEA in 2016. GDP per worker is the total GDP in the grouping divided by total employment. The top 20 MSA are the largest nominal GDP MSAs according to our GDP measure in 2018. Non-MSAs are counties that are not in an MSA, treating each county as a distinct unit.

Figure 6: Inequality within states and state GDP



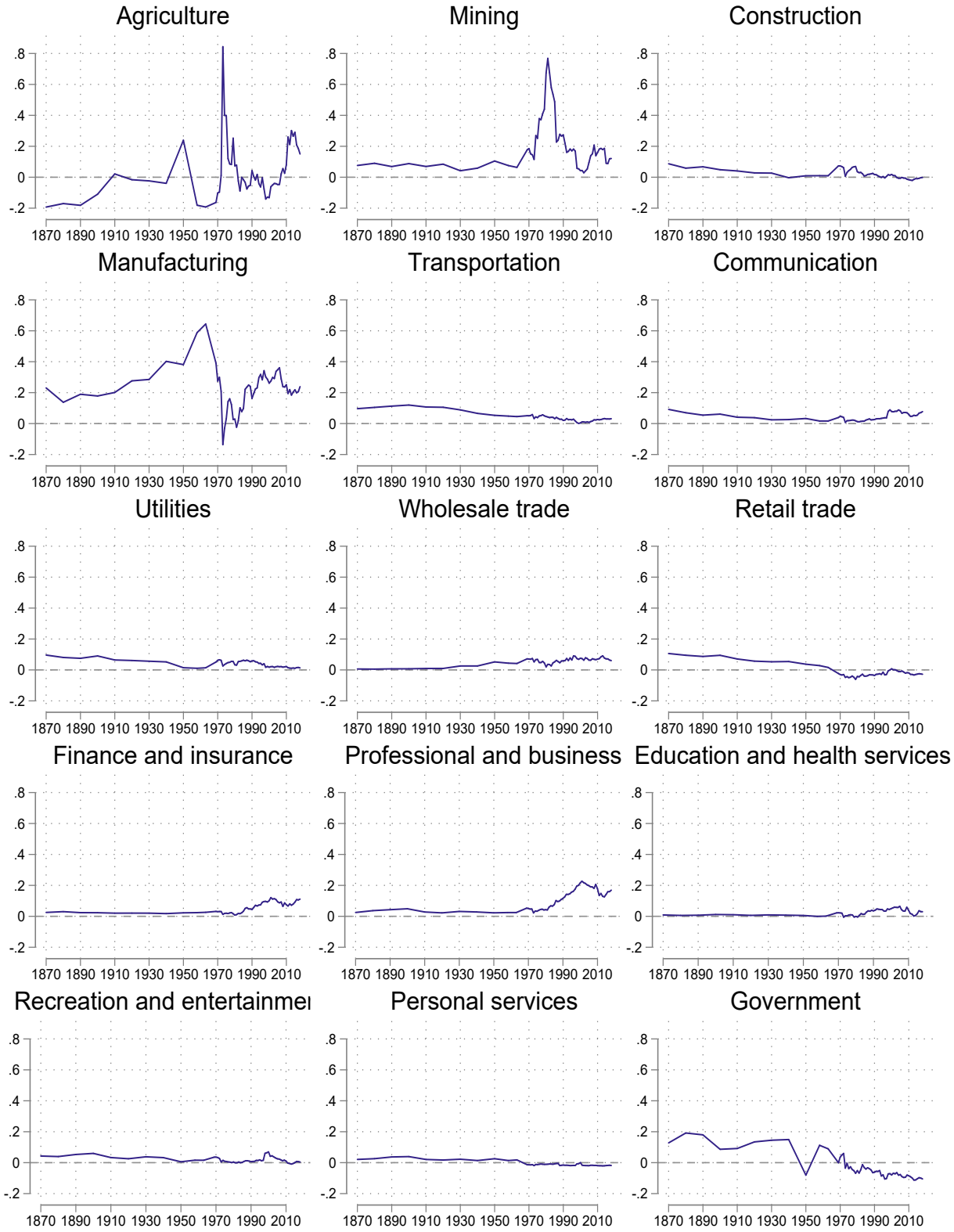
Notes: Each point is a state in a given year showing its (log) GDP per worker calculated by adding counties, and the Theil index of GDP per worker across counties in that state. Lines are the best fit for that year among the states. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 7: Industry share of county GDP by GDP per worker



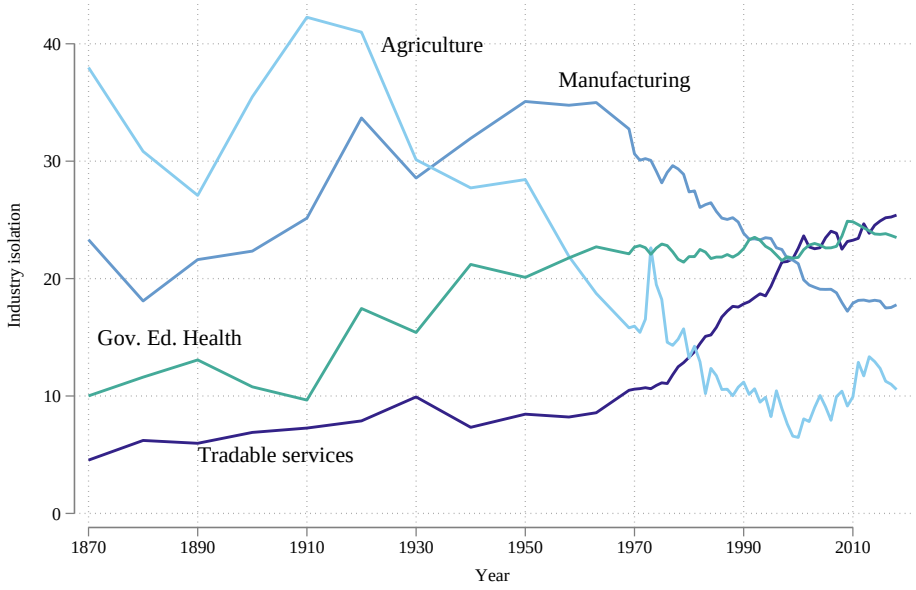
Notes: X-axis is the industry share of county GDP. In each year the top and bottom 0.25 percent of counties in GDP per worker are trimmed. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 8: Sectoral share of GDP per worker inequality as a fraction of the Theil index



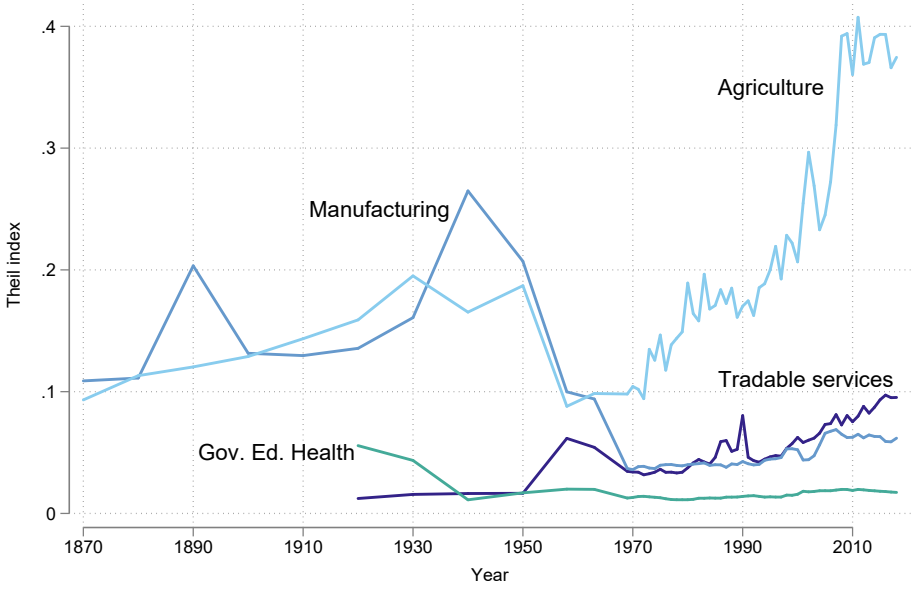
Notes: Real estate not shown. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 9: Industry concentration, isolation index



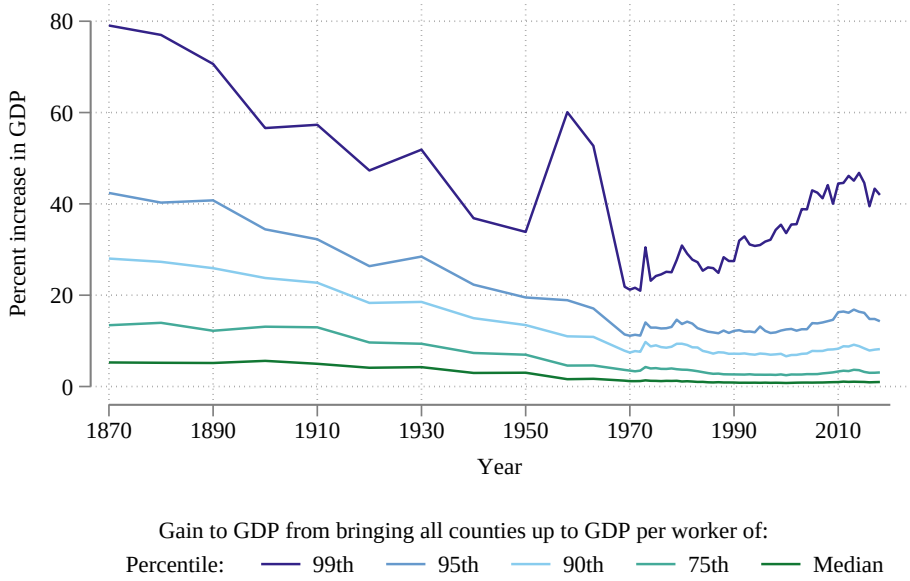
Notes: Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 10: Productivity inequality by industry, Theil index



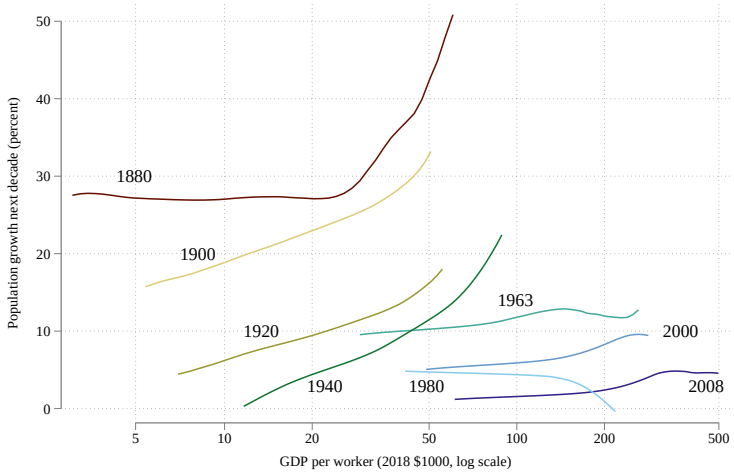
Notes: Productivity county industry GDP per worker. We cannot meaningfully calculate service productivity inequality before 1920, so it is not shown in the figure. We trim the top and bottom 1 percent of the distribution in each year. We note there is a change in data sources in 1969 when the BEA wages series becomes available, so one should be cautious in interpreting changes around that date. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 11: Productivity Gap Index



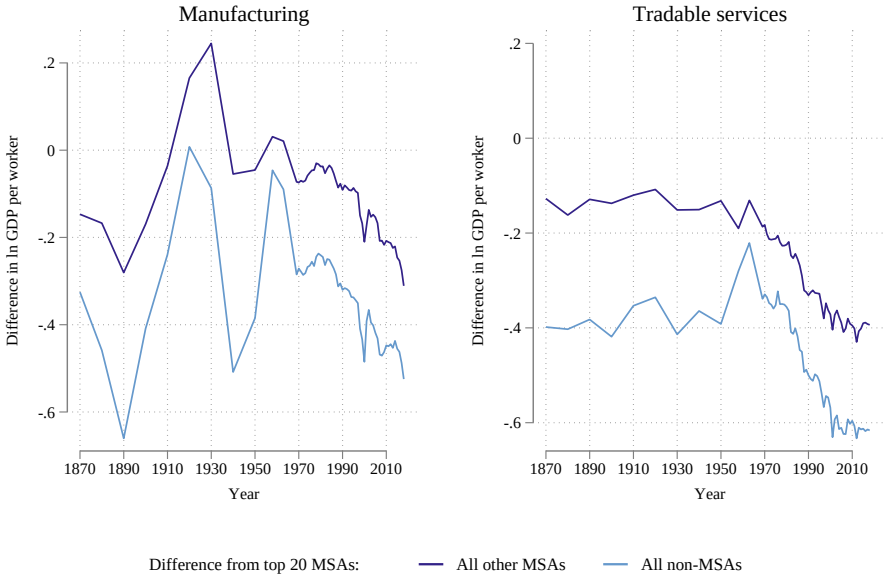
Notes: Shows the Productivity Gap Index, Equation (6), defining the frontier as the counties at different percentile of GDP per worker and the gap as the gain in GDP from bringing counties whose GDP per worker is below the frontier up to the frontier. In caclulating the percentiles, we exclude counties from being in the top percentile that have a mining share greater than 30 percent and have fewer than 10,000 people. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 12: Population growth and GDP per worker



Notes: The lines are local polynomials smoothed. We exclude counties whose absolute value population change was greater than 300 percent and the bottom 5 percent of counties by GDP per worker in a given year. Population growth is over the next decade. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure 13: Manufacturing and Tradable services productivity compared to top metropolitan areas

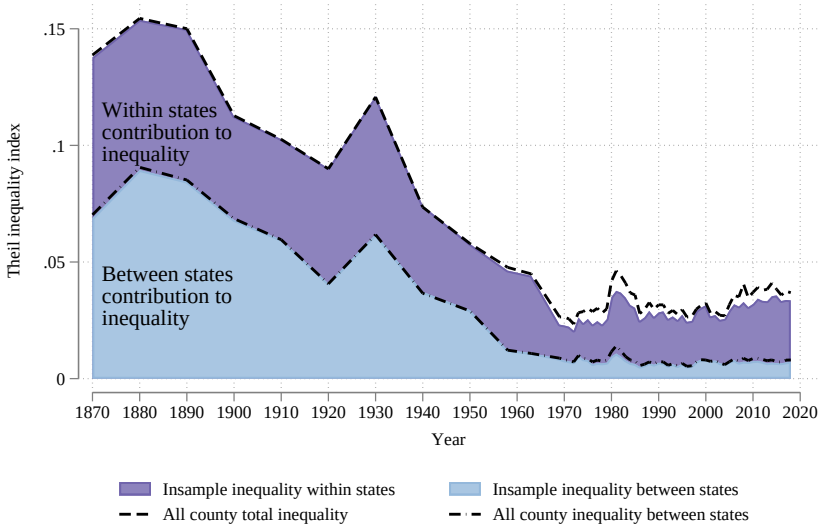


Notes: Each line shows the difference in logs between that group and the top 20 MSAs, so negative values imply that group has a lower productivity than the top MSAs. We use the counties in MSAs as defined by BEA in 2016. GDP per worker is the total industry GDP in the grouping divided by total employment. The top 20 MSA are the largest nominal GDP MSAs according to our GDP measure in 2018. They produce slightly less than 50 percent of all output. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Supplemental Appendix

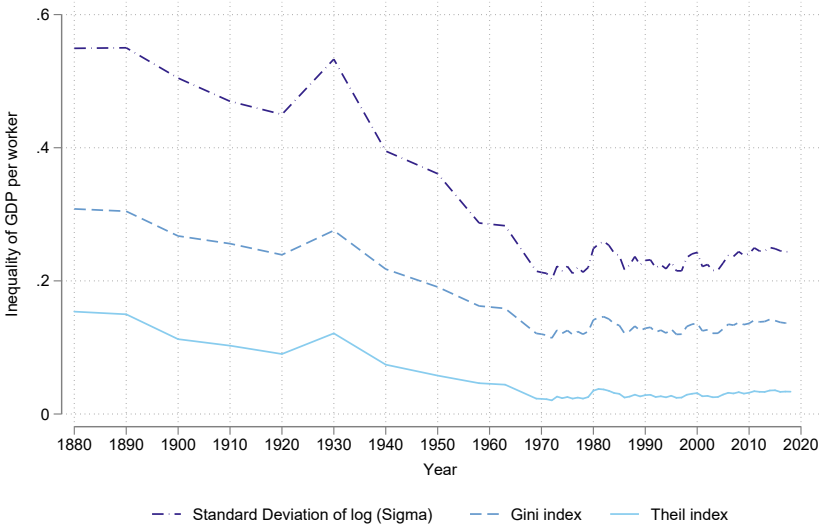
A Additional tables and figures

Figure A-1: Inequality for main sample and all counties



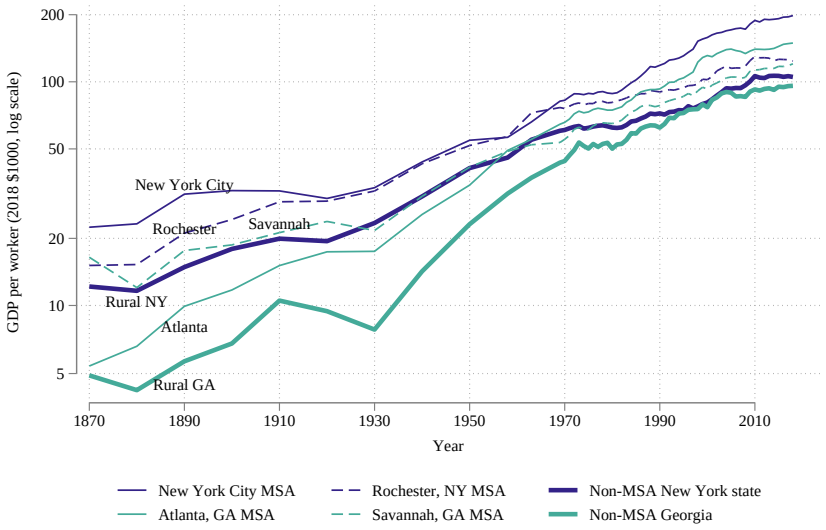
Notes: Compares the Theil index for our main sample (that excludes high mining and utility counties, and very small counties; see Section 2.3) and for all counties, not excluding the high mining and utility counties.

Figure A-2: Comparison of inequality measures for GDP per worker



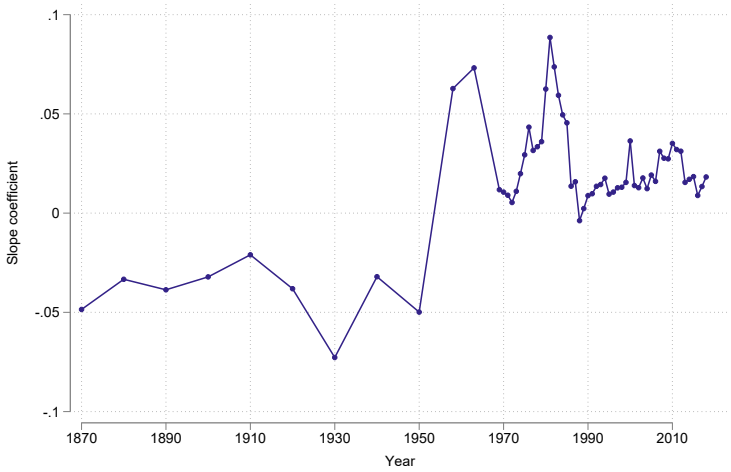
Notes: Shows different measures of inequality of real GDP per person over time. Main sample (excluding high mining and utility counties, see Section 2.3).

Figure A-3: GDP per worker of select metro areas



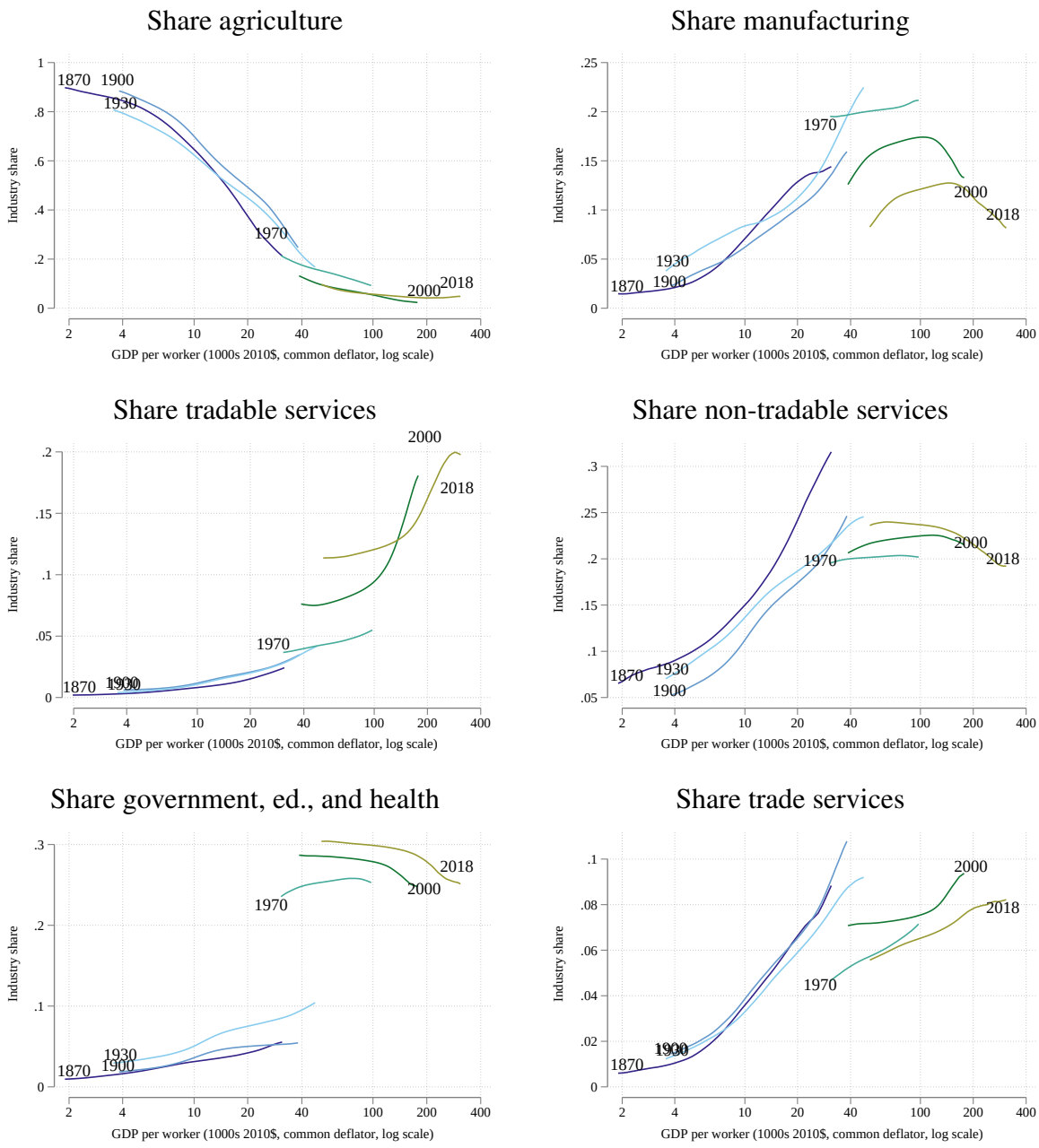
Notes: We use the counties in MSAs as defined by BEA in 2016. GDP per worker (common deflator) is the total GDP in the MSA divided by total employment.

Figure A-4: Inequality within states and state GDP per worker across all years



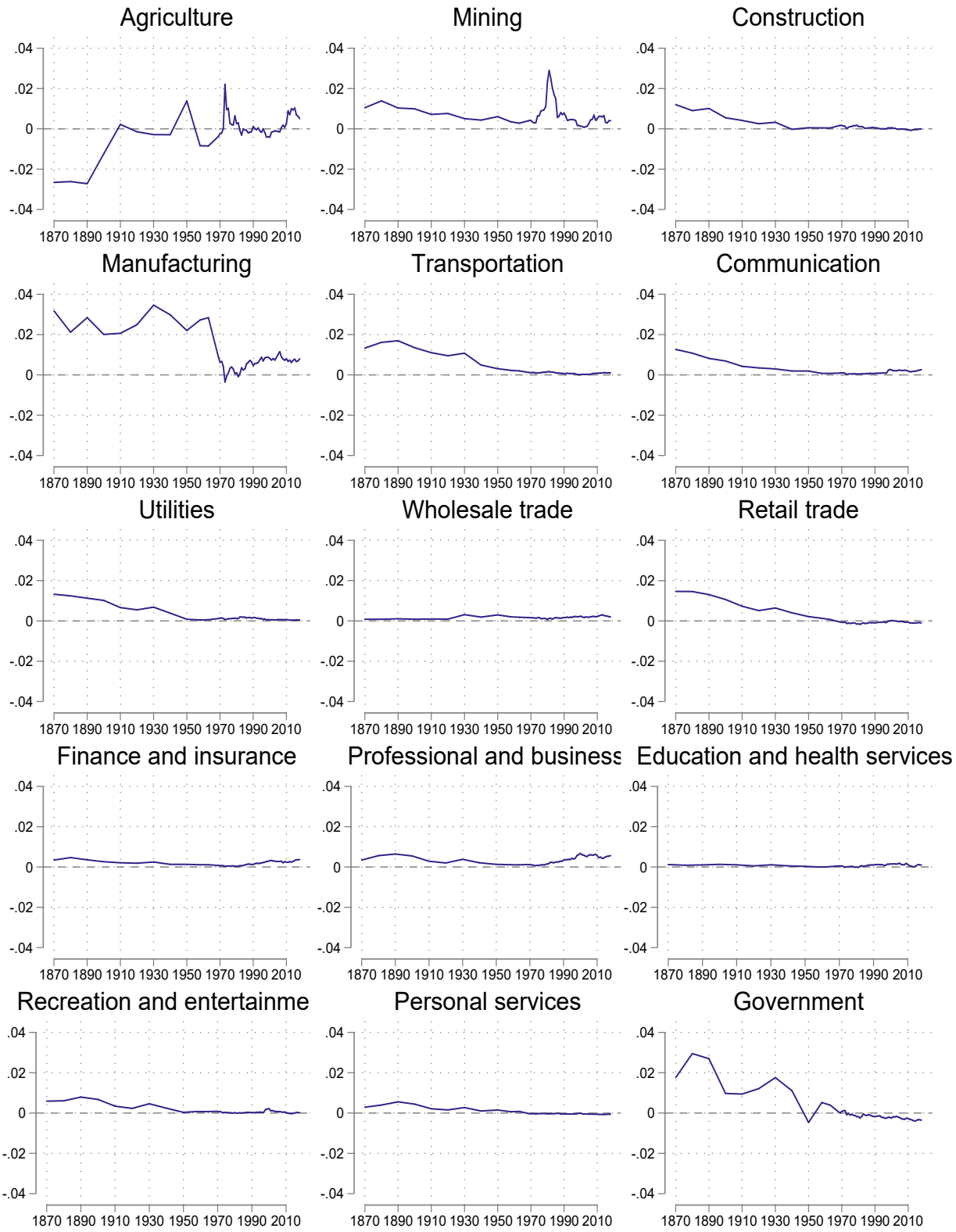
Notes: Each point is the slope of a regression across states in a given year with the dependent variable the Theil index of GDP per worker across counties in each state and the independent variable log state GDP per worker. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure A-5: Employment share by GDP per worker



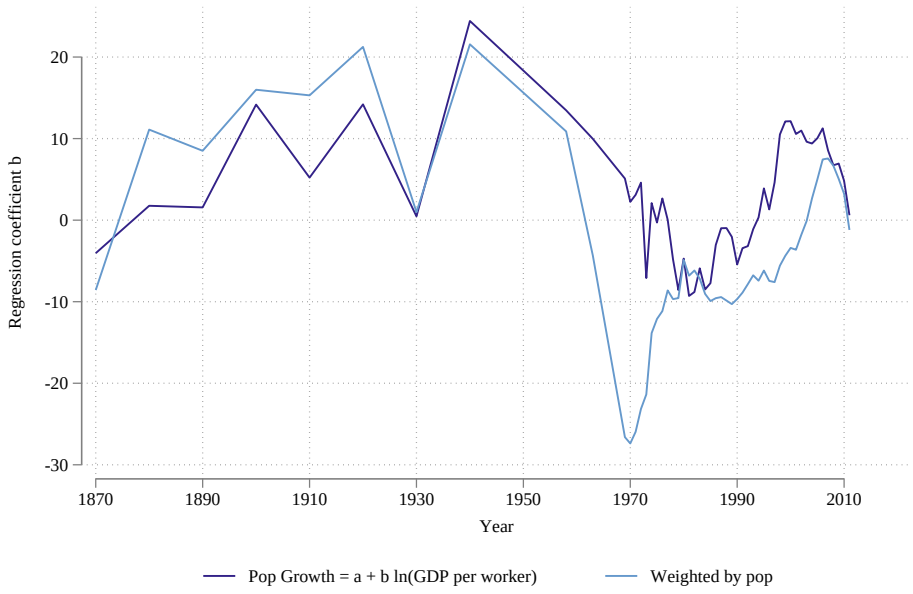
Notes: X-axis is the employment share of total county employment. In each year the top and bottom 0.25 percent of counties in GDP per worker are trimmed. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure A-6: Sectoral contribution to Theil index GDP per worker inequality



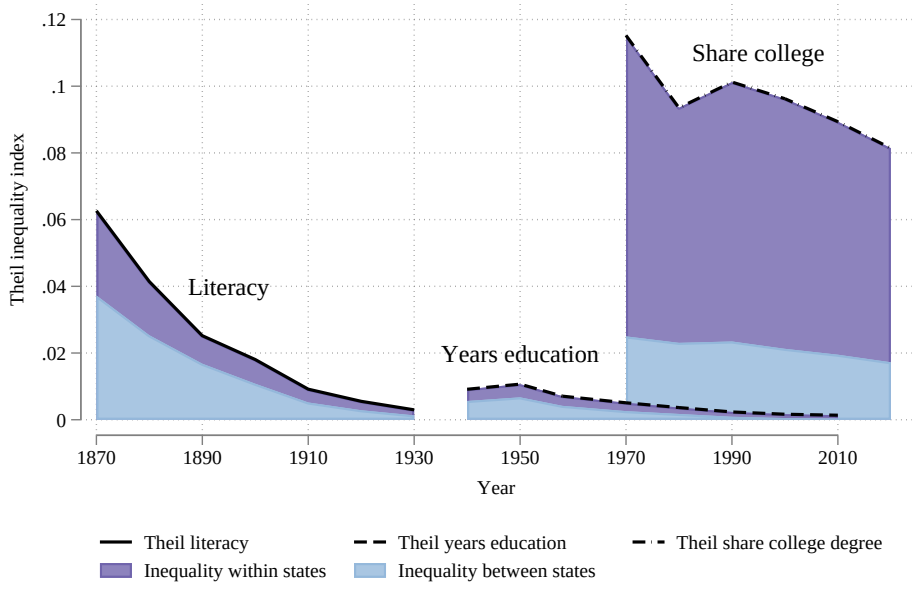
Notes: Real estate not shown. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure A-7: Population growth and GDP per worker regression by year



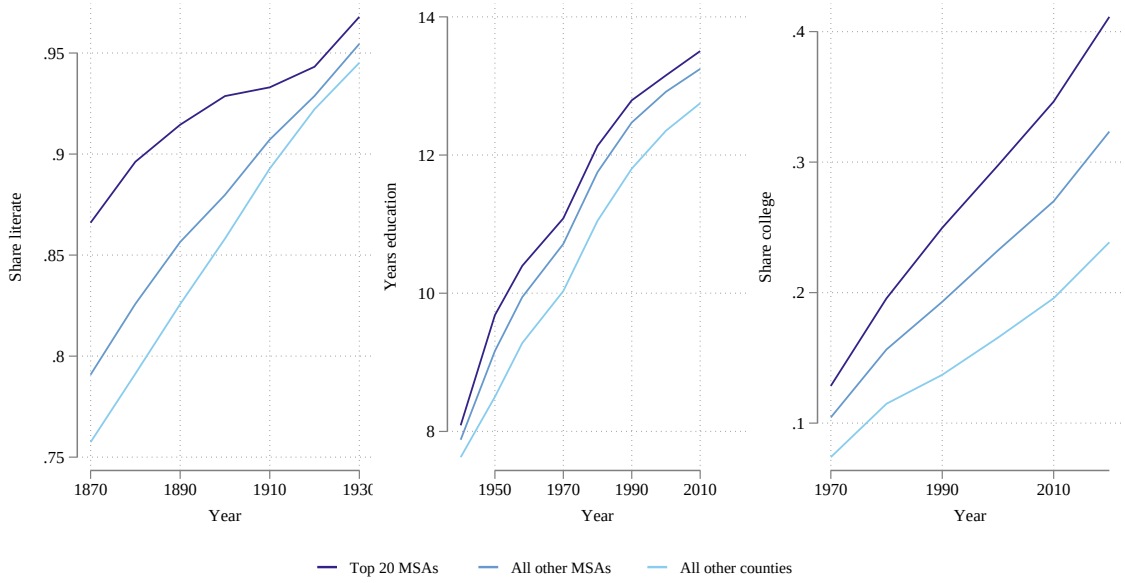
Notes: Each point is the regression coefficient for the linear relationship between Population growth in the next decade and log GDP per worker (common deflator). We exclude counties whose absolute value population change greater was greater than 300 percent and the bottom 5 percent of counties by GDP per worker in a given year. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure A-8: Education inequality



Notes: Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

Figure A-9: Education in metropolitan and non-metropolitan areas



Notes: Averages are calculated using population weight within each group. We use the counties in MSAs as defined by BEA in 2016. GDP per worker is the total GDP in the grouping divided by total employment. The top 20 MSA are the largest nominal GDP MSAs according to our GDP measure in 2018. Main sample (excluding high mining and utility counties, and very small counties; see Section 2.3).

B Data appendix

B.1 Constructing GDP by industry

In this appendix, we describe in detail how we construct GDP estimates for our 16 sectors for different periods. The sectors are 1) Agriculture: Agriculture, forestry, and fishing and agriculture services; 2) Mining; 3) Construction; 4) Manufacturing; 5) Transportation; 6) Communication; 7) Public utilities; 8) Wholesale trade; 9) Retail trade; 10) Finance: Finance and Insurance; 11) Real Estate; 12) Professional: Professional, Scientific, Technical, and Business services; 13) Education and Health: Education, Health, and Social Services; 14) Recreation: Recreation, Arts, Entertainment, and Accommodation; 15) Personal: Personal, Domestic, Repair, and Other; 16) Government.

B.1.1 County GDP 1969-2018

We describe our procedure for each period, starting with the most recent since it requires the fewest assumptions and adjustments and so illustrates the approach. Bureau of Economic Analysis measures of earnings by industry are available at the county level starting in 1969 and state GDP by industry starting in 1963 (U.S. Bureau of Economic Analysis, 2016). We use the earnings by industry to allocate state GDP by industry to the counties within the state using equation (4).

The change from SIC to NAICS classification system in 1997 left these broad industry groupings mostly consistent. The only large inconsistency at this level is the new NAICS Information industry which took some components from several other industries including Arts and Entertainment services and Communications. We included Information in Communications for consistency, although we are unable to fully adjust for the change.

For each year we allocate state GDP by industry to the counties in the state based on the ratio of county earnings by industry to total state earnings by industry. We use the sum of county earnings by industry rather than the total given by the BEA because BEA totals sometimes include allocations not to individual counties. Doing this means all of state GDP will be allocated and so the total per county will add up to state GDP. Very rarely, industry earnings are negative, for example, during a drought in an agricultural county. To deal with this problem in any industry-state-year in which a county had negative earnings, we adjust all counties by adding the absolute value of the earnings of the most negative county to all counties. The total industry earnings are then the sum of all of these adjusted earnings. The most negative county has zero adjusted earnings and so gets allocated zero of state GDP within that industry, while other counties are allocated their proportion of adjusted earnings. For Real Estate in 2008 and 2009 several states have negative earnings in total. We set the state earnings to zero, and so do not allocate any earnings for missing

values to counties. The broader FIRE industry still has positive earnings, and so this procedure allocates GDP entirely based on the other components of FIRE.

Some industries in some counties are not reported to "avoid disclosure of confidential information, but the estimates for this item are included in the [state] totals" according to the county earnings data notes. When a sub-industry is not reported but is part of a larger industry whose total is reported, we allocate the difference between the reported sub-industries and the total based on the state ratios. For example, in the SIC classifications, the total of Transportation, Communication, and Utilities is reported, although the sub-industries Water Transportation and Pipelines may not be. The difference between the total reported for the industry and the sum of each sub-industry gives how much should be in Water Transportation and Pipelines. We divide the non-reported amount based on the state share in the non-reported industries. We follow the same procedure for our fifteen industry groups: we calculate the total not reported for that county, and allocate this amount based on the state share in the not-reported industries. Non-reported industries at the county level are typically small, but matter for small counties since there may often be only one firm in that industry.

B.1.2 County GDP 1850-1940

In this section we describe our construction of county GDP from 1850-1940. The construction is similar to Fulford, Petkov, and Schiantarelli (2016) of county group level GDP, but breaks down GDP into its component parts and makes several improvements. For instance, the calculations use a finer disaggregation of the service sector into eleven sectors rather than six. We use direct calculations of value added for manufacturing and agriculture and construct estimates of value added in the remaining industries using county employment constructed from the full census counts in Ruggles et al. (2019).

County manufacturing and agricultural value added 1850-1940 The census recorded for each county the total value of agricultural output and the value of manufacturing output and costs of inputs. We construct nominal value added of manufacturing by subtracting the cost of inputs from the total output. In 1850, the census did not collect manufacturing inputs. We use the average of the 1860 and 1870 county level ratio of outputs to inputs in manufacturing to calculate manufacturing inputs in 1850. For agriculture during this period the only local measures that exist are of output, not value added. To account for agricultural inputs, we construct a national measure of the ratio of value added to total output by subtracting intermediate inputs from total agricultural output using series K 220 -250 from U.S. Census Bureau (1975). While intermediate inputs were small early on at about 6% in 1850, increasing to nearly 12% by 1900, by 1940 they were nearly 40%. Adjusting for intermediate inputs hastens the relative decline of agriculture after 1900. We apply

the ratio between nominal value added and output at the national level to the value of county level agricultural output to obtain an estimate of agricultural value added at the county level.

The census did not collect manufacturing data in 1910, although estimates of it exist at a national level. To create county level manufacturing, we interpolate between 1900 and 1920 using the national growth in manufacturing value added and allocating growth to each decade so that manufacturing value added grows in each decade in each county at the same rate it does at the national level.

Services, Construction, and Mining employment 1850-1940 We construct county level industry employment using the full individual level census data from the 1850, 1880, 1920, 1930, and 1940; the 5% sample in 1900, and the 1% sample in the remaining years. IPUMS (Ruggles et al., 2019) has created occupation and industry divisions for each individual based on their responses. We allocate the occupations to correspond to the 16 broad industrial categories described in the previous section, roughly following the divisions established by Carson (1949). Doing so creates a measure of the total workers employed in a given industry in each decade. We combine this employment measure with estimates of national value added in each industry allocated to each state and urban or rural area based on the relative wages and employment. We describe the full construction of each of these pieces below, as well as how we combine them using the production theory outlined above.

There are several important difficulties with creating county employment: occupations change over time and some occupations such as legal services that may be classified as a service for an individual are part of manufacturing value added when performed for a manufacturing firm. Since the physical census records from 1890 were largely destroyed by fire, there is no micro-sample from 1890. We linearly interpolate for each county the employment by industry category in 1890 using 1880 and 1900.

Allocating employment by industry We largely follow the mapping of individual occupations and titles into industries used in the census micro data Ruggles et al. (2019) and described in great detail in Bureau of the Census (1950). Although we do not use the employment in farming or manufacturing to construct GDP in these industries, we calculate employment in these industries using the same classification. Military occupations are classified as Federal public administration and so are included in government services.

Value added in the service, construction, and mining industries The construction of value added for services, mining and construction varies by sub-period depending on the information available.

Value added by services category 1840-1900. Gallman and Weiss (1969) construct measures of services value added and employment for eight categories at a national level from 1840 to 1900: trade; transportation and public utilities; finance professional services, personal services, government, education, and "hand trades." Hand trades are composed of smithing, shoe repair, and tailoring. While these occupations may be included in manufacturing output, we include them in the Personal and Repair services for consistency with later periods.

Value added by services category 1930-1960. The National Income and Product Accounts (United States Department of Commerce, 1993) break down the value added by industry from 1929-1958 We use the equivalent tables in United States Department of Commerce (2001) to construct nominal value added in production for the post-war period. The NIPA tables provide only an overall value for Professional, Education, Recreation, and Personal services industries, but provides the compensation of employees in more detailed industries. We divide up the overall value into the four components using the share of compensation in each industry.

Constructing value added for services in 1910 and 1920. No estimates that we have found connect the Gallman and Weiss (1969) and United States Department of Commerce (1993) estimates of services value added by category. To combine these two sets of estimates we assume that the share of each industry in national value added grows linearly between 1900 and 1930. For some services, such as Retail Trade, the shares are almost identical in the two estimates, and so Retail Trade value added grows with nominal GDP. For other services there are larger changes in shares, particularly for Professional Services which Gallman and Weiss (1969) estimates to be a much larger portion of the economy than NIPA, and Finance which is much smaller. Since these two categories have similar value added per worker in 1900, the change has little impact on overall GDP or its county allocation. Finally, since Gallman and Weiss (1969) provides only estimates of the combined Wholesale and Retail Trade, Personal Services (including both Arts and Recreation, and Personal), and Transportation and public utilities (including Transportation, Public Utilities, and Communication), we assume that each component industry grows at the same rate as the combined industry so that their relative contributions to value added in the combined industry are not changing.

Value added for construction and mining. We use the values of mining and contract construction from the National Income and Product Accounts in 1930 and 1940 to construct national value added per worker. From 1880 to 1920 we also use the estimates of Wright (2006) for mining. From 1850 to 1870 we use the ratio of the value added per worker in mining to the value added in transportation in 1880 times the value added per worker in transportation in 1850, 1860, and 1870. This approach assumes that the value added in transportation and mining grew at the same rate from 1850 to 1870. An important part of the value of mineral and fuel extraction comes from transporting it to populated areas. Transportation value added per worker grew at close to the same rate as

overall national product per person during the period. Our approach for construction is similar but involves even stronger assumptions. Construction value added per worker before 1930 is simply its ratio to national income per person in 1930 and 1940. This approach assumes that construction value added grews at the same rate as the national economy, and that employment in construction is a good measure of the distribution of construction activity. Construction is a relatively small component of GDP—it composed only 5% of national product in 1950 and our estimates suggest it was smaller before that—and this approach puts a reasonable value on construction.

Distribution of wages by period and industry For most of the period we only observe wages at the state level, but we use the distribution of wage across counties in 1940 to allow counties within a state to differ. The 1940 census was the first to ask about income or wages and has a complete sample. The sources we use typically do not have coverage for all states in all years. We use the average wage in the census division the state belongs to fill in the wage for missing states in each year. There are nine census divisions which correspond to the broad economic and climatic zones of the United States. Moreover, the state wage for particular occupation can be either a weighted average across urban and non-urban areas or represent an urban or rural wage. We use the state level information and the ratio between urban and rural wages in 1940 to construct estimates of the state-urban/rural specific wages in an industry in the decades up to 1920. We define urban and rural using the 1950 allocation to metropolitan and non-metropolitan areas, the first year such an allocation is available.

1930 and 1940. The 1940 census asks for wages or income, and we use the same classification system for employment to find the average wage within each county and industry. We assume the same geographic distribution of relative wages within each industry applies in 1930.

Wages by state for Trade 1850-1920. We use the wage in each state for bakers (1880-1898 across states, 1907-1928 for select cities) and dress makers (1875–1898 across states) from the United States Department of Labor (1934), page 148 and 219. Since the coverage is fragmentary for different states, we take the average wage over several years to form the decade-by-decade distribution across states. For bakers: 1880 is the average from 1880 to 1887; 1890 the average from 1886 to 1895; 1900 the average from 1891-1898; 1910 the average from 1907-1916; 1920 the average from 1917-1926 we assume the distribution for 1850, 1860, and 1870 follows 1880. For dressmakers: 1880 is the average from 1875 to 1886; 1890 the average from 1886 to 1895; 1900 the average from 1891-1898; we assume the distribution for 1850, 1860, and 1870 follows 1880; and the distribution in 1910 and 1920 follows 1900. Both wages are from urban areas. To form the wage for Trade we take the average of bakers and dress makers.

Wages by state for Transportation. We use the wage in each state for teamsters (male one horse teamsters from 1875-1900 across states, male two horse teamsters from 1913-1928 for select cities)

and engineers (male in locomotive railroad from 1875-1898) from the United States Department of Labor (1934), starting on pages 449, 438 and 453. We convert both series into dollars per day, and we exclude the engineers in states which report only in per mile terms. Since the coverage is fragmentary for different states we take the average wage over several years to form the decade distribution across states. For both occupations we take averages over several years. For teamsters: 1870 is the average of 1875-1880; 1880 the average of 1876-1885; 1890 the average of 1886-1895; 1900 the average of 1891-1900; 1910 the average of 1913-1917; 1920 the average of 1916-1925; we assume the distribution for 1850 and 1860 follows 1870. For engineers: 1870 is the average of 1875-1880; 1880 the average of 1876-1885; 1890 the average of 1886-1895; 1900 the average of 1891-1900; we assume the distribution for 1850 and 1860 follows 1870; and the distribution in 1910 and 1920 follows 1900. We assume both wages are an average from urban and rural areas. To form the wage for Transportation we take the average of teamsters and engineers.

Wages by state for Education. We use the average monthly salaries of teachers in public schools as recorded in the Report of the Commissioner of Education for 1880 (Table 1, Part1, page 408), 1900 (volume 1, Tables 9-10, page 72), and 1915-16 (volume 2, Table 11, page 77). We use the average wage across all male and female teachers, and where the average is not reported for a state compute it as the weighted average of male and female teacher's salaries using the share of male and female teachers in the total. We assume 1850 through 1870 follow the 1880 distribution; 1890 follows the 1900 distribution; and 1910 and 1920 come from the salaries in 1915-1916. These are the average wages for the state.

Wages by state for Mining. We use the wage in each state for coal miners (male coal miners from 1875-1898) and iron miners (male, 1875-1899) from the United States Department of Labor (1934), page 330 and 333. In 1919 we use male hand miners of bituminous coal across states from the United States Bureau of Labor Statistics (1919), Table 3, page 9. Few wages exist early on, and so we use the average of 1880-1889 for the wage distribution in 1880, the average of 1880 to 1889 for 1890, and the average of 1890-1899 for 1900. We assume 1850 through 1870 follow the 1880 distribution; and 1910 and 1920 follows the 1919 distribution. The mining wage is the average of coal and iron mining wages in each year. Both mining wages are for rural areas.

Wages by state for Construction. We use the wages of bricklayers, carpenters, and masons from 1875 to 1928 first for states until 1900 and then select cities from the United States Department of Labor (1934), pages 155, 161, and 190. We assign the city wages to the state and assume that all wages are urban wages. For each occupation we form 1870 using the average of 1875-1880; 1880 average of 1876-1885, 1890 average of 1886-1896; 1900 average from 1891-1900 since the series change in 1901; 1910 average from 1906-1915; and 1920 average 1916-1925. We take the average of the three occupations in each year to form a construction wage.

Wages by state for Government. The Annual Reports of the Postmaster General (1900) recorded

the average compensation of fourth class postmasters by state. It is unclear from the text what frequency the salary is paid, but based on the maximum salary ($4000 to the postmaster general himself), the reported salaries appear to be quarterly. We were unable to find another report that gives a similar breakdown by state. We assume the distribution is the same as 1900 from 1850-1900. We also use the wages of male municipal laborers in sanitation and sewage from 1890-1903 from the Nineteenth Annual Report of the Commissioner of Labor (1905), page 470. We form 1890 using the average of 1890-1895, and 1900 using the average of 1896-1903. We assume that 1850-1880 follows 1890. We form 1850 through 1890 by combining the wages of municipal laborers and postmasters Municipal wages are per hour, and so we combine them with postmasters assuming a 50 hour week and 52 week year. Finally, we form the 1910 and 1920 distribution of wages using the wages paid to police detectives as collected by the Bureau of Municipal Research of Philadelphia (1916). We treat all of these wages as urban wages.

Wages by state for Communication and Miscellaneous Transportation, Professional Services, and Personal Services. We use Transportation for Communication, Education wages for Professional Services, and Trade for Personal Services. These services have a reasonably close approximation to the skill mix in the services for which good wages are difficult to find.

Value added by state for Finance and Insurance. For finance and insurance, we instead use the banking capital by state to allocate national value added. Before 1930, we do not have good observations of wages in finance, and insurance, so rather than follow an earnings allocation approach, it is useful to make some additional assumptions to allocate national value added per worker in finance services based on state banking capital. Suppose capital in financial services is mobile and paid its risk adjusted marginal revenue product across states:

$$MRPK_s = P\alpha A_S K_S^{\alpha-1} L_S^{1-\alpha} = \alpha Y_S / K_S = r_S.$$

Note that here Y_S, denotes nominal value added and K_S real capital. Then, because $Y_{US} = \sum_S Y_S$ and in each state $Y_S = r_S K_S / \alpha$, we can obtain :

$$Y_S = \frac{r_S K_S}{\sum_S r_S K_S} Y_{US} \tag{7}$$

which allocates national value added per worker in finance and insurance to states based on their relative capital returns. If we assume risk adjusted returns are equalized $r_S = r$, which as has some justification, as shown by James (1976), then national valued added can just be allocated by state capital. For 1880--1910 we use the total assets in national banks in each state collected from the annual report of the Comptroller of the Currency by Weber (2000). For 1870, we use the capital of individual banks aggregated to the state level collected by Fulford (2015). We assume that the

distribution of capital in 1850 and 1860 is the same as in 1870 and 1920 is the same as 1910. The 1870 assumption is problematic since the banking capital of the south was largely destroyed by the Civil War, and there were few national banks in the south by 1870. Within states, we allocated state value added to counties using the wage bill or employment, following our standard method.

Combining employment, national value added, and wages We combine the employment, national value added, and wages to create a measure of value added at the county level. In 1940, this procedure is straightforward since we can observe the wage bill and so apply Equation (3) directly using the national value added rather than the state. We similarly use the 1940 wages to apply Equation (3) to allocate 1930 value added using 1930 employment but 1940 wages. This approach assumes that the geographic distribution of relative wages within industry is similar in 1930 and 1940.

We use the relative wages in urban and rural areas in 1940 to create a state by urban-rural distribution of wages before 1930. The exact procedure depends on what kind of wage we observe, but we illustrate with teacher salaries. The teacher salaries we observe are the average for the state. To find the urban and rural wages we take urban and rural education employment in that year and solve for the relative wages that give the average wage so that urban-rural ratio is the same as in 1940.[15]

With a distribution of wages and employment by state and urban, we allocate national industry value added to these larger groupings using Equation (3). Within each state urban/rural area we then allocate the value added using the relative employment of counties using Equation (7).

For Finance, Insurance, and Real Estate, we instead allocate value added to state urban/rural areas using national banking capital in each state and the relative wages in FIRE in 1940. We then use Equation (7) to allocate national value added to state urban and rural areas, and relative employment across counties within each of these areas to allocate value added to counties.

B.1.3 County GDP 1950, 1958, and 1963

Our approach for 1950, 1958, and 1963 is a hybrid of our approach before 1940 and after 1969 based on available data. While it is useful to have the connection from 1940 to 1969, there are many assumptions that vary by year based on what is available. Starting in 1950, the census micro-samples no longer report the current county of residence so it is no longer possible to construct county employment shares by industry. The City and County Databooks (United States Department of Commerce Bureau of the Census, 2012) provide measures of employment, earn-

[15] In equations: we observe the average wage and employment but not the urban and rural wages: $\bar{w}_{s,t} = w^u_{s,t} emp^u_{s,t} + w^r_{s,t} emp^r_{s,t}$. By assuming that $w^u_{s,t}/w^r_{s,t} = w^u_{s,1940}/w^r_{s,1940}$ we have two equations and two unknowns and so can solve for the urban and rural wages.

ings, as well as manufacturing and agricultural products sold. The kind of information they provide changes from year to year, and they do not always perfectly align with what we can construct from the micro-data, so we provide details for each year separately.

Manufacturing. The manufacturing values in the Databooks are reported as value added in 1947, 1954, 1958, and 1963. We use the values of 1958 and 1963 directly since other data match these years and calculate a value for 1950 using the average growth rate in each county 1947 and 1954. Taking the linear average misses the rapid growth during the period, and so we take the average growth rate in each county from 1947 to 1954 and use the county specific growth rate for three years starting in 1947 to find county manufacturing in 1950.

Agriculture. The agriculture values in the Databooks give the total value of farm products sold in 1950, 1959, and 1964. We form 1958 and 1963 values by multiplying the county value by the nominal national increase in the total output in agriculture between 1958 and 1959 and 1963 and 1964 using series K 220-239 in U.S. Census Bureau (1975). We adjust for intermediate inputs and own consumption on the farm using the same series K 220-239. Own consumption was slightly more than 6% of total farm output in 1950. Of much larger importance is the value of intermediate inputs which were close to 40% of total output in 1950.

Mining. The Databooks report a value added measure of mining in 1963, output value in 1954, 1958, and 1963, and employment in 1940 and 1950. However, in 1958 and 1963, the employment, output, and value added of a large portion of counties appear to have been suppressed. These counties had mining reported in 1969 and 1940 but are reported as "Not available" in 1963. Counties with zero are reported as such, so treating "Not available" as zero creates more inequality from mining. We form value added in mining in 1950 by using the 1950 employment and the 1940 distribution of wages. Because of the data suppression, we then interpolate the values in 1958 and 1963 using the national growth rate in mining and the county mining value added in 1950 and 1969.

Construction. The Databooks report county construction employment in 1950 and 1960. We form construction employment in 1958 and 1963 by adjusting for the national growth in the civilian non-agricultural labor force using U.S. Census Bureau (1975) series D11-25. We form value added in construction in 1950 by using the 1950 employment and the 1940 distribution of wages.

Services employment in 1950. In 1950, the Databooks report the employment in transportation and public utilities; wholesale and retail trade; finance, insurance, and real estate; personal services; professional services and overall employment. We use overall employment to construct a residual government employment in 1950 by subtracting out the other categories and setting the residual to zero if it would be negative.

Constructing county GDP in 1950. We combine our value added measures of agriculture and manufacturing at the county level, with valued added in mining, construction, and each of the ser-

vices based on the employment in each of these industries and the 1940 distribution of wages. Our measure of value added comes from Tables 6.1B in United States Department of Commerce (2001) and we allocate the four sub-industries of Services (Professional, Personal, Education/Health, Recreation) based on the total compensation of employees. The 1950 census asked questions about income and wages to only one in five respondents (not all of whom had any income). We therefore use the 1940 wage distribution at the county level for 1950 in equation 4 to form the relative wage, but use the 1950 employment from the Databooks which was originally calculated from the full census. This approach is very similar to our approach before 1940, except that we use actual employment by industry as reported in the Databooks, rather than an estimate based on occupation in the census micro-samples. Since the Databooks do not break down some industries by employment within a county, we use the 1940 county employment ratios in each industry to break Trade employment into Retail and Wholesale; the combined Transportation and Public Utilities employment into Transportation, Communications, and Utilities; and Select Services employment into Education/Health and Professional. Finally, we form the residual of employment not in an industry by subtracting the employment in each industry from the reported total employment. The residual contains Government and Recreation which we allocate based on their 1940 county employment shares.

Constructing county GDP in 1963. In 1963, we allocate state GDP by industry from the BEA to the counties within a state using several measures of local production since we observe employment in some industries, earnings in others, and value added in others. We allocate state GDP for agriculture based on the share of agriculture output in each county. We allocate state GDP for mining and manufacturing based on the share of county value added.

For several industries we observe only employment in 1963. We create a measure of earnings in these industries by calculating the 1969-1970 average earnings per worker in each county-industry and use this earnings distribution to calculate county earnings in 1963. We allocate state GDP for Transportation, Communication, Public Utilities, FIRE, Government, and Construction based on the share of state earnings in each of these industries. We construct government employment as the sum of 1963 local government employment and an estimate of federal government employment based on the ratio of federal to local government in 1967.[16] We form 1963 transportation employment using the 1960 county transportation employment, which includes communication and public utilities in 1963 increased by the national rate of growth in employment. We allocate state GDP in each of these industries based on the overall transportation employment so that counties within a state receive the share of each sub-industry in proportion to their employment in the overall cate-

[16]We construct our estimate of federal employment assuming that local employment and federal employment grow at the same rate in a county from 1963 to 1967: Federal emp. in 1963 = (local government emp. in 1963) * (Federal emp. in 1967)/(local government emp. in 1967).

gory. We construct county employment in Education, FIRE, and Construction using employment in 1960 times the national rate of employment growth. We form professional employment using the share of "white collar occupations" in 1960 and use it to allocate state Professional value added.

Finally, we allocate state GDP for wholesale trade, retail trade, personal, and recreational, based on the earnings in these services reported in the Databooks. We construct Personal and Recreational services using "select services" which include: hotels, personal services, miscellaneous business services, auto repair, repair, motion picture, and recreation. We allocate state GDP in both Personal and Recreational services based on the county earnings in select services and so implicitly assume that the geographic distribution of these services is the same.

Constructing county GDP in 1958. We follow a hybrid approach in 1958, combining the measures of manufacturing, agriculture, and mining value added that we construct for each county from the Databooks with county services constructed by allocating a measure of state GDP based on earnings or employment. We create a measure of state GDP in 1958 by taking the 1963 state GDP in each industry and reducing it at the national rate of growth in that industry based on the NIPA tables in Carter (2006). We construct a measure of county earnings for Construction, FIRE, Government, Transportation, Communication, and Public Utilities using county employment times the 1969-1970 average earnings per worker in these industries. Then we allocate the constructed state GDP for these industries based on their share of total county earnings for each state. We observe only the sum of employment in Transportation, Communication, and Public Utilities, so allocate constructed state GDP for each sub-industry assuming that these industries are in constant employment proportion to each other. We allocate wholesale trade and retail trade based on earnings in these industries; and state personal and professional services based on earnings in select services. We construct Mining using the mining output value in 1958 times the value added-to-output-value ratio in 1963. We allocate state agriculture GDP using state agriculture output compared to the state total. We combine these employment- or earnings-based estimates with the manufacturing value added from the county books, and the agriculture and mining estimates to calculate total GDP in each county.

B.2 Employment by industry

We create estimates of employment by industry from the same sources as our measures of GDP. In many years, constructing employment is an intermediate step to constructing county GDP for many industries and so we describe it in the construction of GDP. Although we use direct value added measures for manufacturing and agriculture, we construct the employment in these industries using the census micro-samples before 1950 since alternative measures are not consistent.

From 1969 on we use the BEA measures of total employment by industry in each county (series

CA25 and CA25N). Because of data limitations, the BEA does not divide between full time and part time jobs, and so the figures are the total number of jobs not the number of people with a job.[17] Since the calculations from the census micro-data before 1950 and from the city and county data books are for individuals, not jobs, the aggregate employment is different between these two series.

We adjust for non-reported or non-disclosure in the same way as for earnings. For codes marked (L) which represent 10 or less, we give 5 jobs. For codes marked (D) which are not reported to avoid confidential disclosure, but which are included in the state totals we allocate the non-reported total proportional to the broader industrial category in a county if it is reported We first calculate the fraction of each industry in a state that is not reported. We then calculate the difference between total employment in a county and employment from the sum of the major industrial categories. Within each county, we then allocate this difference between total and reported employment for the major industries in each county so that it has the same proportion as the not-reported for the state total. This procedure ensures that within each county the allocated total employment or earnings adds up to the reported total.

For the SIC industrial codes (series CA25 from 1969 to 2000), the BEA does not report the employment breakdown within services, or for the combined transportation and public utilities at the county level but does report the breakdown at the state level. We allocate employment within the broader sector using the state earnings per worker within each industry. We observe county earnings in each sub-industry. Variations in sub-industry earnings across counties could come from variations in the amount of employment in a county or from variations in earnings per worker across counties. Since it is implausible that the employment share is constant across counties, we make a more plausible assumption that the relative earnings per worker in each sub-industry is the same across counties, although overall earnings per worker may be higher and lower. We observe the sub-industry earnings $E^j_{c,t}$, and the state earnings $E^j_{s,t}$ and employment $L^j_{s,t}$ in each sub-industry, but not county employment in the sub-industries of services $L^j_{c,t}$. We assume that:

$$\frac{E^j_{c,t}}{L^j_{c,t}} = \phi_{c,t} \frac{E^j_{s,t}}{L^j_{s,t}}$$

for each of the j industries giving j equations for some earnings multiple $\phi_{c,t}$. This assumption implies that the ratio of earnings by janitors and lawyers is the same in New York City as in upstate New York, although both janitors and lawyers may earn more in New York City. Sub-industry

[17] See `https://www.bea.gov/regional/pdf/spi2015.pdf` for a discussion.

labor, which we do not observe, must add up with each county:

$$\sum_j L_{c,t}^j = L_{c,t}.$$

The j earnings per worker equations and sum of labor equation give $j+1$ equations for $j+1$ unknowns (the $L_{c,t}^j$ and $\phi_{c,t}$). And so, it is possible to find the labor breakdown within each sub-industry using the earnings:

$$L_{c,t}^j = E_{c,t}^j \left(\phi_{c,t} \frac{E_{s,t}^j}{L_{s,t}^j} \right)^{-1}$$

and substituting this equation into the total labor equation gives that:

$$\phi_{c,t} = \frac{1}{L_{c,t}} \sum_j E_{c,t}^j \frac{L_{s,t}^j}{E_{s,t}^j}.$$

Place of residence, jobs, and workers When considering production at a fine geographic level, the difference between residence and place of work becomes important. Our measures of employment by industry from the census micro-data before 1950 are by place of residence not place of work. Since commuting across county lines would have been relatively uncommon, the difference between where people lived and where they worked should be relatively unimportant before 1950. In 1950, 1958, and 1963 our measures of earnings and employment come from the City and County Data books. While it is sometimes difficult to tell the definition of the measure of employment in the Data books, the primary source of data is the census and so the earnings or employment location is by residence not place of work. Since we observe agriculture and manufacturing value added in a county directly, these industries are the output of the county. To the extent that commuting is large in these industries, we will tend to overstate the production per worker in some counties and understate it in nearby ones. After 1963, the BEA defines earnings by place of work rather than place of residence.

At the county level, the BEA definition of local employment is by jobs.[18] One person can have multiple part time jobs. The big difference is in self-employment where proprietors can have multiple jobs if they own multiple businesses or are employed but also own a business. This difference creates a large discrepancy between our earlier measures of employment which are based on individuals, not jobs, since the number of business owned by self-employed is much larger than the number of people with jobs. It makes no difference for the allocation of state GDP by county earnings, however, since earnings correctly adjusts for the earnings share of the business.

[18] See p. 76 `https://www.bea.gov/regional/pdf/lapi2015.pdf` for a discussion (accessed 10 April 2017).

The BEA reports in NIPA Tables 6.8B-D the "Persons engaged in production" by industry which is the number of full time equivalent employees and self-employed. To create a consistent series of employment, we adjust the county jobs in each industry so that the sum of the jobs in all counties within each industry is the same as the national "persons engaged in production from the NIPA tables. This adjustment has no effect on the allocation of GDP geographically, instead it shifts the production per worker up by reducing the number of jobs to the number of workers. Adjusting for full-time equivalent by industry has a small effect on relative productivity across counties to the extent that some industries have more self-employed with multiple businesses and industries are concentrated in some areas more than others.

B.3 Creating consistent counties

There are some adjustments that are necessary to calculate consistent geographic measures over more than 160 years. We standardize on 1950 counties using the county boundary definitions from the National Historical Geographic Information System (Minnesota Population Center, 2011). We follow the BEA and early censuses by including Virginia independent cities in their surrounding counties. See `https://www.bea.gov/regional/pdf/FIPSModifications.pdf` for a source of BEA code modifications and the cities that are included in each county.

For historical consistency we also combine counties that have split more recently. We combine the following counties with a single origin county: Menominee, Wisconsin with Shawano; Broomfield, Colorado with Boulder; Ogolala Lakota, South Dakota is Shannon before 2015; Cibola, New Mexico with Valencia; La Paz, Arizona with Yuma.

In addition, since our calculations for services before 1950 rely on samples from the census which are unreliable for small counties, we exclude all counties with a population less than 1000 in all time periods. To mitigate the effect of this exclusion, we combine counties which have a population in 2010 below 1000 with other nearby counties. We combine the next door counties of Grant Hooker, Thomas, Logan, McPherson, and Arthur, Nebraska with each other. We assign the following counties to next door counties, noting the county that was absorbed in capitals: Nebraska BANNER to Kimball; Nebraska HAYES to Frontier; Nebraska BLAIN and LOUP to Custer; Nebraska GARFIELD to Wheeler; Nebraska KEYA PAHA to Brown; Nevada ESMERALDA to Mineral; New Mexico HARDING to Union; North Dakota BILLINGS to Golden Valley; North Dakota SLOPE to Golden Valley; Texas KENEDY to Brooks; Texas BORDEN to Dawson; Texas KENT to Stonewall; Texas KING to Dickens; Texas LOVING to Winkler; Texas MC MULLEN to La Salle; Texas ROBERTS to Hemphill; South Dakota WASHABAUGH has 1020 people and does not show up in BEA to Jackson; South Dakota ARMSTRONG to Stanley; Colorado combine SAN JUAN and MINERAL to Hindsdale; Idaho CLARK to Butte; Montana GOLDEN VALLEY

to Wheatland; Montana PETROLEUM to Fergus; Montana TREASURE to Rosebud.

B.4 County income per person 1950-2014

Starting in 1950, official statistics report measures of personal income per capita at the county level. We combine the county level income data from the County Data Books (United States Department of Commerce Bureau of the Census, 2012) with the county income from the census in 1980, 1990, 2000, and the combined 2008-2012 American Community Survey collected by Minnesota Population Center (2011). In 1950, the census only reported median household income at the county level, while in other years we have mean income per person. To account for this discrepancy we multiply the 1950 median household income by the mean income to median income ratio in 1960 for each county. Starting in 1969, we also use BEA measures of personal income.

C Comparison to other measures

This appendix compares our GDP and GDP deflator measures to other measures.

We begin by showing that the sum our counties closely matches U.S. GDP per capita over time. Figure A-10 shows the sum of our county GDPs divided by their population, and U.S. GDP per person from Sutch (2006). While for services in the early period we use national estimates of value added per worker, modified by local wages, we are not otherwise using national estimates and so the close correspondence is not by construction. Instead, it likely reflects that for agriculture, manufacturing, and services the historical sources that we use are very similar for much of the period to the sources used by Sutch (2006): census estimates of production and employment, augmented in intercensal years by output of commodities such as pig iron that is available at a yearly frequency.. We seem to be systematically lower in the very early years. The census in 1850 did not record occupations for women and was inconsistent for the slave work force before emancipation. We may also be missing output that occurs in counties not yet covered in the census that estimates of early GDP include. In any case, the difference quickly disappears and we match Sutch (2006) nearly exactly after 1890. We take the fact that we match both the level and the growth path well as a strong indication that our overall approach is capturing the broad trends in the economy.

Figure A-12 shows that our approach matches sectoral shares of the economy as well when these become available after 1929, and provide an independent measure of the overall economy before that. Indeed, given the difficulty in calculating shares starting in 1929 and during the subsequent Depression, there is some reason to prefer our more stable estimates in the early years.

Figures A-10 and A-12 also illustrate the massive change in the level and composition of pro-

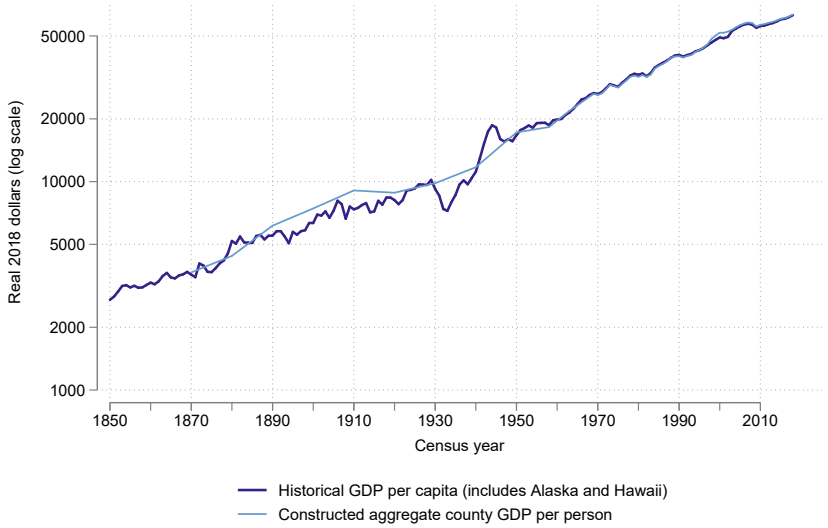
duction that have taken place in the U.S. since 1850. While the U.S. has become much richer, how it produces has shifted, as agriculture declined relative to services as a share of the economy. The share of manufacturing expanded until 1950, but it has since declined. Set against the backdrop of overall growth in Figure A-10, it is important to realize that the decline in manufacturing share in Figure A-12 does not imply a decline in manufacturing production. The U.S. produces more manufactured value added now than in the 1950s, and is much more productive, so it employs fewer people. These aggregate sectoral shifts seem to be part of a broader pattern of development (Herrendorf, Rogerson, and Valentinyi, 2014).

Figure A-11 shows how our county GDP measure compares to state measures of Personal Income from Easterlin (1960), Leven (1925), and Schwartz and Graham (1956). These measures formed the basis for the Barro and Sala-i-Martin (1991) analysis of the convergence across states. Personal Income is distinct from our value-added concept of county GDP because it includes income earned from production done elsewhere and does not include corporate profits. These distinctions suggest that our measure of aggregate state GDP should not be exactly the same as state personal income although the two should be broadly similar. To remove aggregate differences and instead focus on the geographical distribution, we compare each state's share of U.S. GDP based on the county total to the state share of US Personal Income. Our measures closely match the state personal incomes from 1880 to 1940 and by construction match the BEA state GDPs exactly since 1969. The figure shows how different each state's share of total GDP (defined as the sum of county GDPs) is from that state's share of total income. We order them based on the share of total income from that source. Since the shares add up to one, the differences in shares must sum to zero, and so Figure A-11 is essentially graphing the residuals from a regression of state GDP share on state income share. The differences are small, and at least in the early years seem to be systematically related the difference between income and value-added, since the states where we place more GDP than the estimates of personal income are also states with large capital exporting cities.

Finally, the BEA started publishing county GDP estimates for the years after 2001.[19] These estimates use more information than the wage bill, including non-public information on local production. Figure A-13 shows a scatter plot in 2001 and 2019 of log GDP per person calculated using our earnings approach and the new BEA estimates. While most counties align closely, there are a few outliers. The figures show that in both 2001 and 2019, the counties where our measure and the BEA do not closely align are almost all ones where mining or utilities were a large share of the economy. In 2019, the highest GDP per person counties by the BEA estimates are counties where mining is a large share of the economy. Using earnings to allocate mining or utilities may be less accurate because of the high capital involved.

[19] See Aysheshim et al. (2020) for the methodology and `https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas` for the BEA page.

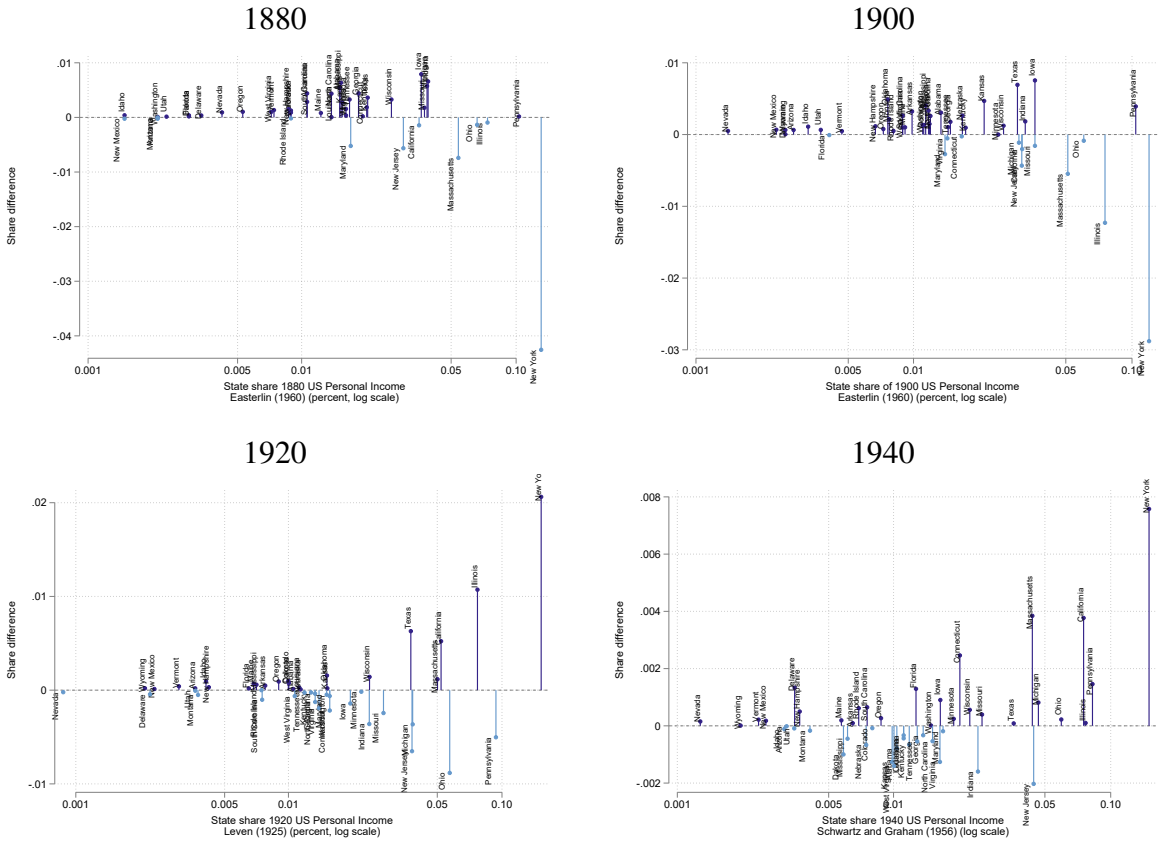
Figure A-10: Aggregate county GDP and estimates of U.S. GDP



Notes: Compares U.S. GDP created by combining all counties with U.S. GDP from Historical Statistics of the United States, Earliest Times to the Present, GDP series by Sutch (2006).

Table A-1 shows simple regressions with our measure of log GDP per person against the BEA measure from 2001–2019. With no exclusions for mining or utilities, our measure is highly correlated with the BEA estimates, with a slope coefficient of 0.829 and and $R^2$of 0.78. Excluding the high mining and utility counties that are the outliers in Figure A-13, however, and the fit improves dramatically. The slope coefficient is statistically indistinguishable from 1, the constant is close to zero and the $R^2$is 0.92.

Figure A-11: Difference between state GDP and state income calculated in other sources



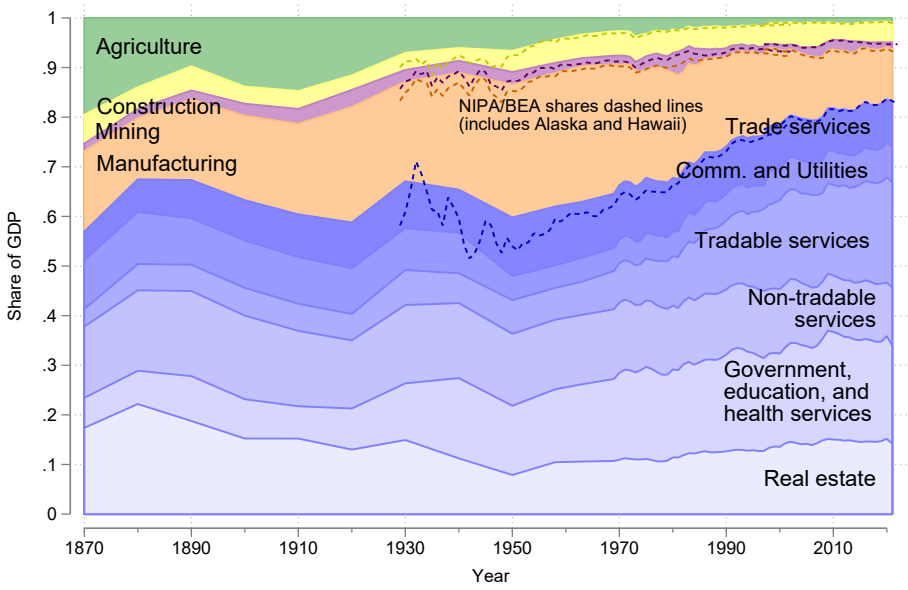
Notes: Each figure shows the difference between the share of total U.S. GDP our county calculations give that state and the share of total U.S. personal income from Easterlin (1960), Schwartz and Graham (1956), and Leven (1925).

Table A-1: Regression comparison of county GDP per person measures

	All counties	In sample	All counties	In sample
	ln GDP per person			
ln BEA GDP	0.829***	0.999***	1.041***	1.053***
per person	(0.00182)	(0.00128)	(0.00178)	(0.00163)
Constant	1.716***	-0.0221*	-0.434***	-0.563***
	(0.0189)	(0.0132)	(0.0188)	(0.0172)
Observations	57,775	53,060	56,080	53,060
R-squared	0.783	0.920	0.949	0.967
Population weighted			Yes	Yes

Notes: Compares the log GDP per person based on our earnings estimates and from BEA from 2001 to 2019. Both measures are deflated using a common GDP deflator. The first column includes all counties for which both measures calculate GDP. The second includes only our main sample that excludes counties with high mining or utilities share of GDP and populations with fewer than 2500. The third and fourth columns weight each county by population size. *** p<0.01, ** p<0.05, * p<0.1

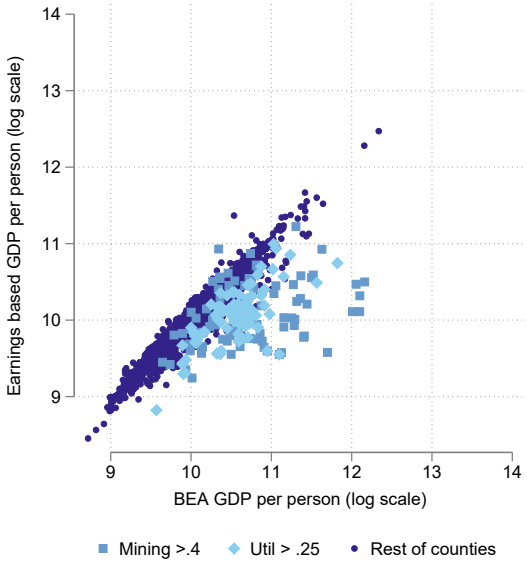
Figure A-12: Industry shares of GDP and employment



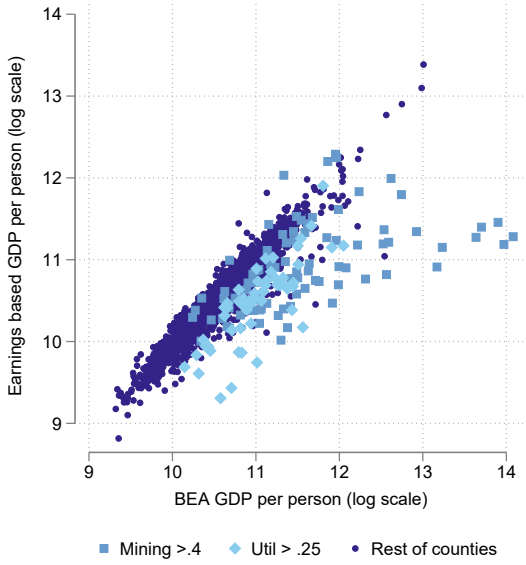
Notes: The solid lines are the implied national share of U.S. GDP by sector from aggregating counties. Industries are grouped by broad sector. The dashed lines show the National Income and Product (NIPA) from 1929–2002 Carter (2006) and 1996–2020 from BEA.

Figure A-13: Comparison of county GDP per person measures

Panel (A): 2001 comparison



Panel (B): 2019 comparison



Notes: The y-axis is the earnings based county GDP per person developed in this paper, the x-axis the BEA county GDP per person estimates, both in a log scale adjusted for inflation with a common deflator. The mining >.4 are counties whose GDP is more than 40 percent mining using the BEA measure, and similarly for utilities.