Working Papers

FEDERAL RESERVE BANK OF BOSTON™

No. 19–20

Recent Employment Growth in Cities, Suburbs, and Rural Communities

Benjamin K. Couillard and Christopher L. Foote

Abstract:

This paper uses a comprehensive source of yearly data to study private-sector labor demand across US counties during the past five decades. Our focus in on how employment levels and earnings relate to population density—that is, how labor markets in rural areas, suburbs, and urban areas have fared relative to one another. Three broad lessons emerge. First, the longstanding suburbanization of employment and population in cities with very dense urban cores essentially stopped in the first decade of the 21st century. For cities with less dense cores, however, the decentralization of employment continues, even as population patterns mimic those of denser areas. Second, a dataset that begins in 1964 clearly shows the decentralization of manufacturing employment away from inner cities that has long been a focus of the urban sociological literature. Starting in the 1990s, however, manufacturing employment fell sharply, not just in cities but also in rural areas, which had experienced less-intense deindustrialization before then. Finally, average earnings dispersion across counties with similar density levels fell during most of our sample period. But after the 1990s, this dispersion rose, probably because of an increase in earnings dispersion among very dense counties ("superstar cities"). We also note that our results are consistent with explanations of rising individual-level earnings inequality within cities that rest on fundamental changes in how basic job tasks are performed, rather than where particular jobs are located.

Keywords: demographics, labor economics, economic geography, location economics, real estate economics, regional economics, spatial econometrics

JEL Classifications: J2, J3, J6, R1, R3

Benjamin K. Couillard is a senior research assistant in the research department at the Federal Reserve Bank of Boston. His email address is Ben.Couillard@bos.frb.org. Christopher L. Foote is a senior economist and policy advisor in the research department at the Federal Reserve Bank of Boston. His email address is Chris.Foote@bos.frb.org.

This version: December, 2019 **https://doi.org/10.29412/res.wp.2019.20**

1 Introduction

> *On top of America's...widening economic divide between the rich and the poor, there is a troubling gap between its geographic winners and losers. The United States is growing spatially more unequal, in ways that are ripping the country apart and threaten to undermine prosperity for all of us.*[1]
>
> —Richard Florida

Economists and policymakers have long studied economic inequality among US households and wage-earners. Lately, this work has broadened to include inequality across places as well. Growing spatial disparities can stem from the interplay between human capital and agglomeration economies. In so-called superstar cities, for example, highly educated workers can exploit agglomeration economies to increase their productivity and earn high wages. At the other end of the spectrum are smaller, distressed areas that are unable to generate adequate economic opportunities for their residents. During the past several decades, many such communities have lost manufacturing jobs that provided a standard-of-living not achievable in other lines of work with similar human capital requirements. Because the residents of distressed cities are often unable to afford the more-expensive rents in more-favored areas, these individuals remain in the distressed areas, so spatial opportunity gaps harden over time.

In October 2019, the Federal Reserve Bank of Boston brought together economists, other social scientists, local leaders, and policymakers to discuss geographic disparities and what might be done about them. This paper informs that discussion by presenting some basic facts about spatial patterns in labor demand during the past five decades. Our main empirical results will draw from a consistent yearly source of private-sector employment and earnings, the County Business Patterns dataset (CBP) constructed by the Census Bureau. The CBP includes consistent yearly data on employment and total private payrolls for all counties in the United States since the mid-1960s, so we can use a historic context to evaluate recent employment patterns in an historical context. The CBP has industry-level detail but does not include individual-level data, so we cannot control for important worker characteristics such as age and education. Additionally, the only industry breakdown we use is the split between manufacturing and nonmanufacturing employers.[2] But the CBP data remains well-

[1]See "America's Polarization Threatens to Undo Us," Citilab blog post, January 25, 2018. Available at https://www.citylab.com/equity/2018/01/americas-polarization-threatens-to-undo-us/551483/. See also the book–length treatment of these issues in Florida (2017).

[2]While higher-level industrial classifications (for example, manufacturing versus nonmanufacturing) are less likely to change over time, the switch from SIC to NAICS in 1998 results in a reduction in the national share of manufacturing employment of 1.8 percentage points greater than indicated for that year by the Bureau of Labor Statistics. See Figure A.10. As discussed below, this is unlikely to affect our overall analysis.

suited to answer questions about when and where spatial developments in labor demand have occurred.

To cite one example, consider the recent job losses in manufacturing. Factory employment has fallen by about one-quarter since 2000, and as an absolute number, this loss of manufacturing jobs is unprecedented in recent US history.[3] Yet as a *share of employment*, manufacturing has been shrinking since the late 1940s; in fact, manufacturing has actually stabilized as a share of total employment since 2010.[4] If the impact of manufacturing on the overall economy depends on its share of employment, as it likely does, then what explains the heightened concern about manufacturing job losses for specific geographic areas today? As it turns out, knowing precisely when and how manufacturing employment reallocated across different areas in the United States since the early 1960s helps us answer this question. The CBP's data on annual payrolls is also useful for linking developments in individual-level inequality to broader spatial patterns. For example, labor economists have long known that the return to completing college began to rise in the early 1980s and levelled off in the 2000s. As we will see, it is striking how these time-series patterns also emerge in county-wide earnings data as well.

Our study of the CBP generates three broad lessons. The first concerns the suburbanization of employment and population, which was a common feature of cities with very dense urban cores in the late 20th century (Glaeser and Kahn 2001, 2004). Consistent with some previous research, we find that in the last decade or so, the suburbanization of both population and employment essentially stopped for cities with dense cores. The decentralization of employment for large cities with somewhat less-dense cores continues, however, even as their population patterns mimic those of denser areas. The re-centralization of population in both types of major US cities is consistent with the increased attractiveness of urban life that has been suggested by previous work (Guerrieri, Hartley, and Hurst 2013; Couture et al. 2019). In fact, that continued movement of *employment* to the suburbs outside of the handful of densest cities may indicate that centers of cities have become more attractive not because it is better to work there (perhaps because of increased agglomeration economies) but instead because it is better to live there (because of better amenities).[5]

A second finding of the paper is that changes in the spatial pattern of manufacturing employment over the past five decades could help explain the seemingly outsized effects

[3]According to the Bureau of Labor Statistics, manufacturing employment averaged 17.3 million over the 12 months of 2000, but had fallen to 12.7 million by 2018. The latter number is slightly higher than the trough of 11.5 million manufacturing jobs reached in 2010.

[4]Manufacturing employment was 32.1% of total nonfarm employment in 1953 and 8.5% in 2018.

[5]In addition, the comprehensive nature of the CBP allows us to study decentralization in smaller cities. For those areas, decentralization has never been an important feature of the data. This is most likely because congestion in smaller and rural cities has always been modest, so there was no pressure to move population or employment to outlying areas.

of manufacturing declines on communities today. Data from 1964 onward shows clearly that manufacturing employment decentralized markedly away from inner cities and toward rural areas in the 1960s and 1970s. This movement of factories out of inner cities has of course been a focus of the urban sociological literature for some time, most notably in the work of William Julius Wilson (Wilson 1987, 1996). Starting in the 1990s, however, manufacturing employment fell sharply not just in cities, but also in rural areas, which had experienced less-intense deindustrialization before then. To be sure, manufacturing as a share of employment has declined in both urban and rural areas during the past 50 years. But this decline was much less rapid in rural areas. Moreover, the CBP data indicate that, holding density constant, areas with substantial manufacturing employment were somewhat insulated from the overall manufacturing decline before the 1990s. Another way of saying the same thing is that in the first part of our sample, manufacturing employment became increasingly concentrated in areas that already had a lot of manufacturing employment. In the 1990s and 2000s, however, deindustrialization reached these manufacturing centers as well. This change in the spatial pattern of deindustrialization may help explain why manufacturing declines seemed to have more severe effects in the 1990s and 2000s than during earlier decades.[6]

Finally, we present some results on inequality of earnings both across and within cities. For most of our sample period, average earnings dispersion across counties with similar density levels fell over time. But starting in the mid-2000s, this dispersion rose because of a large increase in earnings dispersion among the densest counties. Such movements would be consistent with the relatively recent emergence of dense superstar cities, which have begun to pull away from less-favored cities that have the same densities. Additionally, our results are consistent with explanations of rising *within-city* inequality of individual-level earnings that rest on fundamental changes in how basic job tasks are performed in the United States today, rather than where particular jobs are located.

The remainder of our paper is organized as follows. Section 2 describes the CBP dataset and the modifications we make to it. This section also outlines the four broad density-based classes of counties that allow us to study rural areas, suburbs, and dense cities within a common framework. Section 3 presents some basic data on the spatial distribution of employment and population in counties of varying densities, and section 4 delves further into the spatial distribution of manufacturing employment. Section 5 discusses the distribution of earnings, which we measure as total private-sector payrolls paid in the county divided by the total number of jobs. Sections 6 and 7 relate our findings to current research on the effects of manufacturing declines in different areas and the increasing income inequality

[6]As we discuss below, the spatial explanations for manufacturing-shock effects that we explore are complementary to explanations offered by Eriksson et al. (2019), who explain localized effects of recent manufacturing declines using a product-cycle model.

within large cities, respectively. Section 8 concludes.

2 Data from County Business Patterns (CBP)

2.1 Constructing a Balanced Dataset

For county-level data on employment and earnings from 1964 to 2016, we turn to County Business Patterns (CBP), an annual extension of the Economic Census that relies on administrative data from the Census Bureau's Business Register as well as survey data.[7] CBP reports employment, payrolls, and establishment counts by size class. Although employment and payrolls are sometimes suppressed to prevent disclosure of the operations of a single firm, establishment counts never are. The advantages of CBP over other county-level datasets from Bureau of Labor Statistics (BLS) and Bureau of Economic Analysis (BEA) include the CBP's detailed industrial breakdown and a more complete coverage of educational, membership, and small nonprofit organizations. The disadvantages of the CBP are that it excludes public administration, crop and animal production, the postal service, railroad employment, certain financial firms, self-employed individuals, and private households. Additionally, like the alternative county-level datasets from the BLS and the BEA, the CBP has no individual- or worker-level data.[8]

As discussed below our regression methodology is enhanced when the data are balanced, so we make a number of modifications to the CBP data, all of which are detailed in the appendix. Briefly, we combine a small number of counties into county groups in order to replicate county borders as of 2019.[9] Other combinations of counties are warranted by their extremely small size, which can generate outsized yearly *percentage* changes in employment from time to time.[10] We also combine independent cities in Virginia (which are official county equivalents) with their surrounding counties. In grouping counties for any reason, we never group a county with one in another state. Finally, we exclude all counties from Alaska and Hawaii due to their distance from the rest of the country. All told, our algorithms reduce the total number of counties and county-equivalents in the continental United States from 3,108 to 2,909, and we refer to this group of 2,909 as "counties" for the remainder of the

[7]The 1964–1969 data were supplied by Ody and Hubbard (2011) to the Inter-University Consortium for Political and Social Research (ICPSR), where we obtained the data. ICPSR is also our source for the data from 1970–1973, which were placed there by the Census Bureau. The 1974–1985 data come from the National Archives, and the 1986–2016 data were downloaded from the Census Bureau.

[8]For more information on CBP, as well as our imputation strategy for suppressed data and our procedure to ensure a balanced panel, see the appendix.

[9]For example, Broomfield County, CO, was created in 2001. We create a fictitious Broomfield County going back to the start of our dataset in 1964 by combining appropriate fractions of employment and payrolls from the surrounding counties from which Broomfield was created. See the appendix for details.

[10]Our rule is to combine counties with less that 200 employees with neighboring counties.

paper.

2.2 Classifying Counties on the Basis of Density

Our initial look at the CBP data groups counties into four groups based on population density in a given year, as illustrated by Table 1. Because a county's population density can change over time, counties can move among these groups over time. The first density group includes the least-dense 85 percent of US counties, and Table 1 shows, these counties account for about 30% of the US population in any given year (more precise population shares are given below). The next group is made up of counties in percentiles 86—95 and accounts for about one-quarter of the US population, while Group 3 includes percentiles 96–99 and includes about 30% of US residents. The final group is comprised by the densest 1% of counties. In any given year, the densest 29 counties ($\approx 0.01 \times 2,909$) make up about 15% of the national population.

Table 1 also shows that on average, at least 30% of counties in the lower three classes border a county in a denser class. For counties in density percentiles 86–95, the share is around one-half, which illustrates that many of these counties are suburbs for denser cities. Figure 1 illustrates this point with a map of the US counties in 2016 that is based on the four density-group definitions. The two densest groups are depicted by shades of blue, while the two least dense groups are in shades of red. The map makes clear that urban areas typically consist of a dense center county (or groups of counties) and outlying counties of declining density. Figure 2 makes this point more clearly by focusing on the Northeast Census Region in the years that begin and end the sample period (1964 and 2016). Boston, New York, and Philadelphia always include at least one county in the densest 1% of counties, and in New York's case, several central counties are so designated. The presence of counties in each of the four density classifications in most of the nation's largest urban areas—and the relative stability of county-level classifications over time—suggests to us that our grouping scheme is useful for analyzing long-run labor market trends.

A comparison of the two maps for the Northeast illustrates, however, that counties can and do change density groups over time. In 1964 Pittsburgh's center county (Allegheny County) was among the densest 1% in the nation, but this county had dropped to the next-highest group by 2016. Table 2 lists the 29 densest counties for five selected years: 1964, 1973, 1989, 2000, and 2016.[11] Only 35 counties are ever included included among the densest 29 counties in any of the five years. Consistent with the general migration of population toward the Sun Belt, counties in the metro areas of Los Angeles (Orange County), Tampa–St. Petersburg (Pinellas), Atlanta (DeKalb), Dallas, and Northern Virginia join the top

[11] As we will later use these years to measure changes in various statistics over time, we have chosen the three middle years (1973, 1989, and 2000) because they are roughly at business-cycle peaks.

1% during the sample period.[12] Perhaps surprisingly, Los Angeles County never appears in this group, although its neighbor Orange County does. Los Angeles County extends from the dense urban area near the coast into the sparsely populated area around Angeles National Forest, while Orange County is more uniformly urban. In our analysis below, we have ensured that our substantive results are robust to the inclusion of Los Angeles County in the top 1%.

3 Employment, Population, and Decentralization

This section outlines some broad trends in population and employment for counties in the different density groups. We then discuss trends in suburbanization and decentralization among counties within the same labor market area.

3.1 County-Level Population and Employment

The top two panels of Figure 3 depict shares of population and population growth for the four density groups. The bars in the top left graph show shares of population in the groups for the five selected years. As noted in Table 1, these shares range from around 30% for Groups 1 and 3 to less than 20% for the densest-group (Group 4). Over time, the largest (and offsetting) changes in population shares occur in Groups 2 and 4. Most notably, the population share of the densest counties declines in the first half of the sample period and stabilizes thereafter.

Because the specific counties included in the four density groups can change, we also examine the population-density relationship in a different way. In the top right panel of Figure 3, the bars represent annualized population growth by initial density group and year. For example, the uppermost bar in the panel shows the average (unweighted) population growth between 1964 and 1973 for those counties that started out in the least-dense group, regardless of whether those counties moved into the next dense group during those nine years. The two most striking features of this graph appear in the top and bottom bars, which correspond to the least-dense and most-dense group of counties, respectively. The top bars show that for the rural counties, there was a significant step-down in population growth between 2000 and 2016. Before then, rural growth rates had been similar to the growth rates among the other groups. Indeed, the corresponding bars in top left panel show that rural share of population ticked up somewhat between 1964 to 2000. In the new century, however, the rural population share declines, consistent with the significant decline

[12] Note that the Northern Virginia group of counties, which includes Fairfax and Arlington counties, is an example of a grouping we made when adjusting the CBP sample.

in population growth among rural counties after 2000 depicted in the top right panel. The opposite story is told for the most dense counties. The panel at left shows that population growth among very dense counties, which had been negative from 1964 through 1989, swung positive after that year. Correspondingly (and as noted earlier), the panel at right shows that the population share of the densest counties stabilizes in the last half of the sample period.

The two bottom panels display the same information for employment rather than population, but the disparate outcomes for the least-dense and the most-dense counties late in the sample period are broadly similar. Among rural counties, employment growth declines sharply from 2000 to 2016; declines for the other groups are not as sharp. It is also true that the panel at right shows that for the densest counties, employment growth falls to near zero from 2000 to 2016. But our investigation of the data indicates that this weak employment growth is driven largely by the experiences of the two declining counties, which drop out of the densest group between 2000 and 2016: Cuyahoga County, OH (in Cleveland) and Orleans Parish, LA (in New Orleans). As seen in Table 2, their spots in the top group are taken by Dallas County, TX and DeKalb County, GA (in Atlanta). When these new counties are included in the densest group, the *share* of population accounted for by the nation's densest counties remains stable, as shown in the panel at bottom left. This stability in the dense-county employment share is consistent with the stabilizing share of *population* in the densest counties in the top left panel noted earlier. And the relatively good fortunes of the densest counties on the employment front stand in contrast to the employment record of the most rural counties, which see their share of employment decline after 2000. All told, the four panels of Figure 3 are consistent with the growing concern that the social and economic fortunes in the most-rural and most-urban parts of America have bifurcated during the first two decades of the 21st century.[13]

3.2 Decentralization and Suburbanization in Commuting Zones

We now turn to localization of population and employment within wider labor market areas. In recent work, commuting zones (CZs) are typically used to aggregate counties into distinct labor market areas. These zones, which are based in large part on commuting patterns, are mutually exclusive and include all counties in the country.[14] Given the recent gains in very dense counties, a natural question is whether dense counties are gaining relative to other

[13]For a discussion of the economic and social challenges facing rural America, "The Hard Truths of Trying to 'Save' the Rural Economy," by Eduardo Porter, *New York Times*, Dec. 14, 2018. Available at https://www.nytimes.com/interactive/2018/12/14/opinion/rural-america-trump-decline.html.

[14]Commuting zones were constructed by the Economic Research Service of the US Department of Agriculture, and we use the 1990 version of these zones. For more information on how commuting zones are defined, see https://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas/.

counties in their CZs. Such centralization would stand in contrast to the trend towards suburbanization and decentralization in metropolitan areas during the late 20th century (Glaeser and Kahn 2001, 2004).

To investigate this issue, we assemble our counties into CZs as defined by the government in 1990.[15] We then classify each CZ according to the density group of its densest county. Table 3 lists the CZs that include a county in the densest 1% in each of the selected years. With some abuse of terminology, we will call this list the "dense CZs," even though these CZs may include outlying counties that are not dense at all. In all but one year (1973), 17 CZs meet this definition, but only 13 CZs are in this category in each of the five years. Cincinnati, Cleveland, New Orleans, and Pittsburgh fall out of the list over time, while the Sun Belt cities of Atlanta, Dallas, Los Angeles, and Tampa Bay enter it.[16]

The changing membership in the list of dense CZs matters if one is concerned about population shares for the dense CZs over time. Defining groups on the basis of 1964 characteristics will impart a downward bias to population growth of dense CZs over time, because the group would include declining CZs like Cleveland and Pittsburgh. Conversely, defining the group on the basis of 2016 characteristics will overestimate population growth among the dense CZs, because the group will then include the growing Sun Belt cities.[17] As it turns out, however, the timing of the definitions does not matter for our question of interest: whether activity is centralizing or decentralizing within CZs over time. The top panel of Figure 4 depicts the average share of CZ population that is accounted for by the CZ's densest county, using the 1964 groupings of counties. The dark blue line shows that for CZs with very dense cores, this line trended steadily downward from the start of our sample to the early 2000s, consistent with the decentralization and suburbanization over this period noted by Glaeser and Kahn (2001). In 2006, however, the line flattens out. The lighter blue line corresponds to CZs whose densest county is in the 96th through 99th percentile of county density in 1964. Here again there is a trend toward decentralization that flattens out in the mid-2000s. The tan and red lines show that among more-rural CZs with less-dense cores, there has never been a trend toward decentralization, perhaps because the relative sparsity of core counties in these CZs generated little pressure for it. The lower panel of the figure

[15] Our groupings of some counties generally respected commuting zone borders, so it is not difficult to group our counties into these zones. However, a few very small commuting zones had to be combined in order for this to occur. As a result, although there are 722 1990-era commuting zones in the continental United States, there are only 712 CZs in our sample. See the appendix for details. An alternative to using CZs would be to use the government's definitions of Core Based Statistical Areas (CBSAs). Unlike CZs, however, CBSAs do not include all counties in the country, just the metropolitan and "micropolitan" areas.

[16] Note that even though there are always 29 counties in the densest 1% of counties, there are fewer than 29 CZs with a very dense county. This is because some CZs (particularly New York) include more than one of the 29 counties in the densest 1%.

[17] See Table A.2 in the appendix for a graph of population growth by CZ density group using these two definitions.

shows the same information using the 2016 classification of counties. Although there is some change in the levels of the lines across the two panels, the shapes of the lines are similar. For the two groups of CZs with the densest core counties, the longstanding trends toward suburbanization of population appear to have slowed or stopped in the mid-2000s.

Figure 5 measures the centralization of employment.[18] As with population, the timing of our group definitions does not seem to matter much, as the shapes of the lines in the two panels are similar. And also similar to the population results, the suburbanization of CZs with the densest cores appears to have slowed in the early 2000s. Perhaps the most important difference between the population and employment figures concerns CZs with core counties in the 96th–99th density percentiles. The lighter blue lines in Figure 5 provide less evidence for a slowing of the employment decentralization in the mid-2000s compared to the evidence for a slowing population in decentralization in Figure 4.

A critical question now facing urban economists is why big cities appear more attractive to potential residents today than they did in the previous decades. Is the increased attractiveness due to rising benefits of cities in facilitating *consumption*, particularly of luxury amenities that may not available in less-dense areas (Glaeser, Kolko, and Saiz 2001; Couture et al. 2019)? Or are bigger cities thriving now because of increased agglomeration effects in *production* (Baum-Snow, Freedman, and Pavan 2018)? The results presented in this section obviously cannot distinguish between these two stories, but it is clear that something important about big cities changed as the 21st century began.

4 Spatial Evolution in Manufacturing Employment

4.1 The Urban-to-Rural Migration in Factory Jobs

Another key question among regional and urban economists is why the ongoing declines in manufacturing employment appear to have larger adverse effects on communities today in than in the past (Charles, Hurst, and Schwartz 2018). The top two figures of Figure 6 present shares and growth rates of manufacturing employment, in the same way that this information was depicted for total employment in the lower row of Figure 3. Because our CBP data go back to 1964, the figure is able to capture a significant feature of US manufacturing employment during the last decades of the 20th century—its migration from urban to rural areas, even as manufacturing became a smaller part of employment at all density levels.

The top left panel of Figure 6 shows the *share of national manufacturing employment*

[18] The jump in the line corresponding to the 1% densest counties between 1973-74 is the result of a change in methodology by the CBP. This change is small relative to the total decentralization in the entire sample period. See the appendix for more information.

in each density group. The two least-dense groups account for a steadily growing share of national manufacturing employment from 1964 to 2016, with a naturally offsetting movement in the national share experienced by the two denser groups. These trends are particularly pronounced for the densest and least-dense groups. The bottom panel shows the *share of manufacturing out of total county-level employment* for each group. The three densest groups experience declines of about 30 percentage points, while the least dense group sees a decline of only about 15 percentage points. In short, manufacturing becomes more rural during our sample period, even as all of the density groups come to depend less on factory jobs. The top right panel depicts the annualized growth rates of manufacturing employment for each group and adds some color to the story. Even though the rural share of manufacturing continues to rise after 2000, the top right panel shows that more-rural counties experience very large factory-job losses in percentage terms for the first time in our sample period.

4.2 A Regression Model of Manufacturing Shares

We can make more precise statements about the spatial evolution of manufacturing employment with a regression model that relates local manufacturing shares not only to population density, but also to the general migration of the US population toward the Sun Belt and to local education levels. Consider a series of yearly cross-sectional regressions specified as

$$\begin{aligned} \mathbf{Y_t} &= \alpha_t + \beta_t \mathbf{X_t} + \gamma_t W \mathbf{X_t} + \varepsilon_t, \\ \varepsilon_t &= \lambda_t W \varepsilon_t + \nu_t. \end{aligned} \tag{1}$$

Here, $\mathbf{Y_t}$ is an $N \times 1$ vector of county-level manufacturing shares[19] in year t, $\mathbf{X_t}$ is an $N \times K$ matrix of regressors, W is an $N \times N$ spatial weighting matrix (which is constant across years), and ε_t is an $N \times 1$ vector of unobserved errors.[20] The second equation indicates that we are estimating a series of *spatial error models*, because we are allowing the errors of nearby counties to be correlated with one another. In this context, "nearby" is defined by the spatial weighting matrix W, which we define as a second-order contiguity matrix, and the time-varying strength of the correlation is captured by the estimate of λ_t.[21] As is common,

[19]While there is a drop in the mean of this variable that is 1.8 percentage points greater than would be expected in 1998, there is no drastic change in the behavior of any coefficient after this point, with the exception of the indicator for a county being in the 5% densest. However, this could very well be mostly or partly the result of a change in the composition of this indicator; the 5% densest vary by year and there is a similar jump in that coefficient in the early 1980s. Furthermore, this coefficient in this regression does not contribute to our central argument.

[20]To clarify some notation: Because the constant term is the same for all counties in a given year t, α_t is single scalar multiplied by an $N \times 1$ vector of ones: $(\alpha \cdot \iota_N)_t$. The errors in ε_t, however, vary across county-level observations for a given year t.

[21]Before normalization, a second-order contiguity matrix places a 1 in the (i,j)'th entry when county i shares a border with county j. If county j shares a border with some other country that also shares a border

we row-normalize the weighting matrix by dividing all the entries in the i'th row by the sum of all the entries in this row. In this way, the entries in each row of the normalized matrix sum to 1, so when applied to either the regressors $\mathbf{X}$ or to the errors ε, W generates weighted averages of neighboring-county values.[22]

The list of potential regressors $\mathbf{X}$ includes

- The natural log of the county's population density in year t,
- A dummy variable denoting whether the county is among the densest 5% of counties in year t,
- Average January temperature of the county calculated from 1941 to 1970, and
- The natural log of the share of county population with bachelor's degrees.[23]

To facilitate the interpretation of results, each of these variables is standardized, by year, to have a mean of zero and a variance of 1. The inclusion of $W\mathbf{X}$ in the model implies that we can also include some nearby values of regressors. Formally these are known as *spatial lags*. Although we experimented with including spatial lags for most of the regressors in the list above, only the spatial lag for log population density tended to be significant, so only that lag is included in the baseline model.

Figure 7 displays maps of county-level manufacturing shares for the first and last years of the sample period. The map for 1964 shows that manufacturing was particularly important for counties in the eastern half of the country in the mid-1960s. By 2016, manufacturing had declined sharply throughout the United States, with particularly large losses in the Northeast. Figure 8 takes a first pass at explaining these declines with a regression that includes the three density-related variables (plotted across the top row of panels) and January temperature (plotted in the bottom left panel). The λ_t coefficients that capture the intensity of the spatial-error correlation are depicted in the bottom middle panel. Consistent with the bar charts in Figure 6, the top row of panels shows that manufacturing employment generally rotated to less-dense counties during the sample period. The decline in the linear

with county i, then the (i,j)'th entry in W equals 0.5. See Figure A.3 in the appendix for an illustration of the second-order contiguity matrix.

[22]Using spatial methods are critical when estimating county-level models. If one ignores the positive spatial correlation that is likely to exist across counties, then the resulting standard errors will be severely understated. Conditional on choosing an appropriate form of the weighting matrix W, spatial models generate the right standard errors and impart efficiency gains as well. There is a parallel between estimating a spatial model in our context and estimating a Prais-Winsten or Cochrane-Orcutt correction when errors are correlated across time rather than space. As long as the time-series process for errors is specified accurately, then the time-series corrections will generate appropriate standard errors as well as more-efficient estimates.

[23]We construct a yearly measure by interpolating values from the 1950, 1970, 1980, 1990, and 2000 Censuses, as well as the 2008-12 and 2013-17 ACS 5-year averages.

log-density term shown the top left panel is particularly steep, although these coefficients level off in the mid-1990s.

Figure 9 adds county-level shares of residents with bachelor's degrees to the model. The partial correlation of this variable with manufacturing share declines until around 1995, indicating that manufacturing employment becomes relatively more prevalent in counties with lower education levels until the mid-1990s. At that point, the coefficient estimates reverse course. Given the large absolute declines in US manufacturing employment in the 2000s, this reversal indicates the factory-job losses in less-educated areas were more severe.

The inclusion of an education variable in the model also has a significant effect on the log-density coefficients. As seen in the top left panel, these coefficients continue to decline over the sample period, indicating that manufacturing shares in less dense counties few grew steadily relative to the shares of denser counties. But these movements are less pronounced in the model that includes education as compared to the earlier model that omitted it. Moreover, there is no longer a significant inflection point in the density coefficients in the mid-1990s. The implication is that the movement of manufacturing to rural areas was not driven completely by lower density *per se*, but also because manufacturers found it more advantageous to locate in less-educated areas.

A common feature of the models with and without education is the behavior of the spatial error terms, λ_t. In both models, the strength of this correlation is very high at the start of the sample period—near 0.80—indicating that the typical regression error for a given county is about 80% of the weighted average of the errors for surrounding counties. Like the residuals in any regression, errors in this model reflect omitted variables that determine the dependent variable. For the location of manufacturing employment, these variables might include historical influences, easy access to raw materials or particular customer markets, local government policies, etc—each of which is likely to be spatially correlated at the county level. What's particularly interesting about the errors in both Figures 8 and 9 is that the degree of spatial correlation drops sharply in 1995—about the same time that the education coefficients rise. This pattern indicates the "clusters" of counties with strong manufacturing presences began to be adversely affected to a much greater degree after the mid-1990s than before. We return to this point below.

5 The Spatial Distribution of Earnings

5.1 Nonparametric Analysis of Earnings and Density

The inclusion of total payrolls in the CBP data also allows us to study the evolving relationship between population density and average earnings. We define average earnings as a county's total annual payroll divided by total employment, that is, total earnings per job. Because the CBP does not include individual-level data, we cannot tell precisely why average earnings in a county may have changed. Its residents may have become better educated, they may have become older and thus more experienced, or employment in the county may have shifted to higher-paying industries and occupations.[24]

A simple and nonparametric way to analyze earnings and density is with binned scatter plots, which are presented in Figure 10. Each of the six panels corresponds to a selected year. Each dot in the panels depicts the mean of total payrolls for a given percentage of county-level density; thus there are 100 dots in each panel. This mean is taken relative to a population-weighted mean of average payrolls across all counties in the given year. Because the US population is concentrated in very dense counties, and because earnings-per-job is a positive function of density, the vast majority of less-dense counties lie below the zero line in each panel.

Figure 10 reveals an interesting evolution in the wage–density relationship over the past half-century. In 1964, there is a "crooked smile" relationship between average earnings and county-level density. The minimum for average earnings appears to be somewhere near the 30th percentile of density, and the earnings–density relationship rises gradually for counties with higher densities after that point. Over time, however, the crooked smile becomes a hockey stick. By 2016, there is little relationship between density and average earnings for the lower two-thirds of the density-percentile distribution. Then the relationship begins to rise, with a substantial increase in the slope of this relationship among the very densest counties. Earlier, we saw that dense counties were doing better in terms of population and employment growth, and that CZs with very dense cores were no longer decentralizing to the same extent as in previous decades. Here, we see that dense counties are doing better in terms of earnings as well.

Of course, the fact that the *average* dense county is doing better in terms of earnings does not mean that *all* dense counties are doing better. To see how average earnings are distributed within counties with the same population density, the top two panels of Figure 11 depict the 75th and 25th percentiles of earnings per job within each percentile. (The vertical dotted lines in these panels are explained below.) The top left panel shows that in 1964, the gap between the 75th and 25th percentiles—also known as the interquartile range (IQR)—was fairly similar across the density distribution, although there is some narrowing of the IQR at the far right, among the densest counties. The top right panel, however, shows

[24] The CBP includes industry-level detail on employment and establishment counts, but the suppression of data becomes more common as industry-level disaggregation increases. Total payrolls for a county are virtually never suppressed, but we found it quite difficult to impute payrolls at the industry level—even payrolls in manufacturing, where employment can be imputed based on the size distribution of establishment counts. We have therefore left an industry-level analysis of earnings-per-job at the county level for future work.

less narrowing of the IQR as county density increases in 2016.

The lower three panels of Figure 11 present yearly averages of IQRs across three segments of the county-density distribution—the bottom 85% of counties, percentiles 86–95, and the top 5%. These segments are demarcated in the top two panels by vertical dashed lines. The blue lines in each of the lower panels display the IQRs without weighting the counties by population, while the red lines are the averages of population-weighted IQRs.[25] Comparing the three lower panels, it is clear that at the start of the sample, the average IQRs for the least dense counties are higher than the density among the top 5%; for example, in 1964 the average unweighted IQR among the bottom 85% of counties is about 0.25, while the corresponding IQR for the top 5% is less than 0.20. Thereafter, the average IQRs for all three segments remain fairly stable or decline, indicating a trend toward less dispersion in average earnings conditional on county density. But in the late 1990s and early 2000s, average IQRs rise for each of the three segments. The increase in the average IQR among the top 5% of counties, depicted in the bottom right panel, is especially strong.

As with the results in the previous section involving growth and centralization of population and employment, the results for average earnings add support the broad contention that urban and rural labor markets are pulling apart. Among the top 5% densest counties—where close to half of the country lives—there is a steep gradient of average earnings with respect to density. Moreover, income divergence among the densest counties suggests that a new class of high-earning cities is pulling away from the rest.

5.2 Regression Models of Earnings Levels

We can also use a spatial regression model to relate average earnings not only to density, but also to county-level education levels and to local manufacturing shares. The regression model in this section will include the same regressors those in the model in the previous section, although manufacturing share (the previous dependent variable) will be added to the list of regressors, along with its spatial lag. Figure 12 illustrates this spatial lag by depicting the 1980 county-level manufacturing shares of employment in the top panel and spatial lag of 1980 shares in the bottom panel. Comparing the two panels makes clear that the matrix W essentially creates weighted averages of nearby variables. The lower panel also drives home the migration of manufacturing activity to southern states, consistent with the urban–to–rural migration noted in section 3 and the spatial migration-share regressions above.

Figure 13 presents the results of the average-earnings regressions, which suggest that the years around 1980 were critical for the changing geography of earnings. The top left panel

[25] The data in the top two panels of Figure 11 are not population-weighted.

graphs parameter estimates for the log of population density (solid line) and the spatial lag of this variable (dashed line). As we would expect, the solid line shows that own-county density always has a positive effect on earnings; people in urban areas tend to earn more than people in rural areas. But the size of the coefficient on the log density variable begins to decline around 1980. Around the same time, the coefficients for the spatial lag of density rise temporarily, although these coefficients are again insignificant by the end of the sample period. The top right panel displays the estimated coefficients on the dummy variable indicating that the county is among the top 5% in terms of density. This coefficient begins to rise around 1980, becoming significant around 1990. Taken together, these two panels are broadly consistent with the nonparametric analysis of earnings and density in Figure 10. In particular, the relationship between density and the top 5% dummy variable captures the emergence of the hockey-stick shape among the binned scatter plots late in the sample period.

In the middle row, the left panel shows a rising effect of January temperature on county-wide average earnings during the late the 1960s and the 1970s, which stalls out around 1980. Interestingly, this pattern is consistent with the well known result that per capita income convergence across US *states*—a feature of the data for most of the 20th century—stalled in the 1980s as well. The middle left panel shows the bachelor's degree coefficients, and the path here is similar to the behavior the individual-level college premium. This premium dipped somewhat in the 1970s, rose sharply during the 1990s, and leveled off in the 2000s. The main difference between the coefficients from the county-level model and the individual-level college premium concern the degree to which the coefficients rise after the early 1980s. The middle left panel shows that they rise only to their level in the mid-1960s, but the individual-level college premium rose far above its mid-1960s value. Notably, this coefficient reaches its nadir in 1980.

The bottom left panel depicts the coefficients on own-county manufacturing share[26] (solid line) and the spatial lag of this variable (dashed line). The solid line shows that having a large manufacturing presence in a county tends to generate high wages. This could be a mechanical effect arising from the wage premium that has traditionally been paid to manufacturing workers. The own-manufacturing effect declines over time, yet remains significantly positive at the end of the sample period. The dashed line shows that, conditional on a county's own manufacturing share, the effect of a large manufacturing share in neighboring counties is consistently negative. One interpretation of this pattern is that manufacturing has tended to be located in areas with lower wages, conditional on the other variables in the model. Note also that the neighboring-share coefficient declines somewhat in the 1970s, but reverses

[26]Despite the jump in this variable in the aggregate in Figure A.10, there is no jump in this coefficient in 1998.

direction in 1980. This pattern is consistent with what has been termed "domestic offshoring" of manufacturing employment toward lower-wage areas within the United States in the 1970s and 1980s, and then a substantial loss of manufacturing jobs in low-wage areas after that.

Finally, the lower right panel depicts the estimated spatial error terms, λ_t. These coefficients trend downward until the early 2000s and rise thereafter, consistent with the decline in dispersion in earnings seen nonparametrically in Figure 10. In the previous section, earnings dispersion was calculated conditional on density; that is, dispersion was calculated among counties in the same density percentile across the nation. Here, the spatial error term measures the *spatial component* of earnings dispersion conditional on all of the variables of the model. A high value of λ_t therefore suggests that there are groupings of nearby counties that have high (or low) average earnings conditional on all the variables in the model, including the density variables.

As it turns out, the density measures are the key inputs that drive the precise down-and-up pattern of spatial error terms. To see this, we estimated the model several times, including a limited number of variables each time. As seen in Figure 14, if no covariates are included, then the λ_ts declines somewhat in the first half of the sample period, but is roughly constant after the mid-1980s. A model with only the log density term, however, does a much better job in approximating the full-model λs at least through 1995.[27] Summing up, both the earlier nonparametric analysis in Figure 10 and the parametric spatial model in this section point to density as the most important conditioning variable when analyzing dispersion in county-level earnings. Further, both of these analyses also indicate that this conditional variance declined for most of the sample period, but increased in the 2000s. And the nonparametric analysis suggests that this increase is driven in large part by counties in the top 5% of the density distribution.

6 Consequences of Recent Manufacturing Declines

6.1 Employment-Growth Regressions

The next two sections use our data to shed light on two policy-relevant topics. This section addresses a puzzle: although manufacturing has declined as a share of employment for virtually the entire postwar era, the negative effects of this decline on local communities appears to have been greater during the past few decades (Charles, Hurst, and Schwartz 2018). A local economy's ability to diversify and its level of human capital are no doubt critical in determining its ability to withstand manufacturing job losses. Our data indicate

[27] When the 5% density dummy is included along with the linear density term, the resulting spatial error terms more closely approximate the terms from the full model. The additional effect is small, but grows over time, consistent with the earnings distribution becoming more like a hockey stick, as shown in Figure 10.

that recent factory-job losses have been experienced by communities that are probably less able diversify and have lower education levels, so the consequences of recent declines in manufacturing have been greater. In the next section, we discuss a second issue: what trends in earnings inequality *within* cities should mean for policies designed to reduce geographic disparities.

To address the manufacturing puzzle, we estimate a series of spatial error models that project near-term employment growth onto the manufacturing share of a county and on the average manufacturing shares in neighboring counties. The dependent variable for these models is employment growth from year t to $t+3$.[28] We build intuition for manufacturing's effect on overall employment growth by first estimating two separate models for growth in manufacturing and nonmanufacturing employment, respectively. The right-hand-side variables for both models are the same as those for the average-earnings regressions in the previous section: in addition to own-county and neighboring-county manufacturing shares, we also include the three density-related variables (log population density, neighboring log density, and the 5% dummy variable), college shares, and January temperature.

In the appendix, we present all of the coefficients from these models.[29] Here we discuss the most important coefficients: those for own-county and neighboring-county manufacturing shares. The top two panels of Figure 15 display these coefficients from the manufacturing-employment regression.[30] The solid line indicates that for virtually the entire sample period—and especially early on—those counties with high manufacturing shares in year t saw less *growth* in manufacturing employment between years t and $t+3$. The implied effect is large, as the estimates for the first few years of the sample period are around 0.18. Dividing by three to give a rough estimate of the 1-year effect, and noting that the manufacturing-share regressor is standardized to have unit variance, indicates that a one-standard-deviation increase in a county's own manufacturing share reduces the subsequent 1-year growth rate of factory employment by about 6 percentage points.[31] However, the dotted line in the same panel shows that at least until the 1990s, this negative impact is offset to a large extent if the county is embedded among other counties that also have high manufacturing shares.

[28] That is, the regressors are dated as of year t, but employment growth was calculated over the ensuing three years. Using 3-year rather than 1-year growth rates made the coefficients easier to interpret, as they were less subject to sampling error. However, the results using 1-year rather than 3-year growth rates were qualitatively similar, albeit noisier.

[29] See appendix Figures A.4, A.5, and A.6 for the full results from the employment-growth regressions.

[30] As in previous exhibits, the own-county coefficients are depicted by the solid line and the spatial-lag coefficients are shown by the dotted line.

[31] Here, the jump in aggregate share manufacturing (see Figure A.10) in CBP does matter, causing measurement error in both the right- and left-hand-side variables. We would expect measurement error in the 1995-1998, 1996-1999, and 1997-2000 growth rates to bias the coefficients towards zero, and these coefficients are substantially higher than those in the surrounding years.

The top right panel adds the own-county and neighboring-county effects together by showing the total effect[32] of a one-standard-deviation increase in both types of shares.[33] Although this total effect is negative for most years, it is substantially smaller than the own-county effect in the top left panel.

The middle left panel of Figure 15 shows the same coefficients from the regression with nonmanufacturing employment on the left-hand side. As we would expect, these coefficients here are much smaller in absolute value. A large own-county manufacturing share raises nonmanufacturing employment growth in the county somewhat, especially early in the sample. The neighboring-county manufacturing-share effect is also relatively small, except for the two recessionary periods during the late 1970s and early 1980s, when it is strongly negative. The middle right panel shows that the two effects add up to a positive effect of initial manufacturing shares on nonmanufacturing employment until the mid-1990s (again, aside for the recessionary periods).

Finally, the bottom row of panel shows the effect of the manufacturing shares on overall employment growth. The coefficients in this row therefore combine the manufacturing and nonmanufacturing effects estimated separately in the top and middle rows. The bottom left panel shows that on the whole, the very large manufacturing effects shown in the top row dominate the overall employment-growth effects, even though less than half of the employees in a county work typically work in manufacturing plants. All told, a large own-county manufacturing share tends to reduce near-term overall employment growth for most of the sample period. However, early in the sample, this negative impact was offset to a large degree by the effects of a large *neighboring* manufacturing share. And the top row shows that this pattern was driven in large part by the effect of the initial manufacturing shares on subsequent manufacturing employment. How should we interpret these results? The most important message is that early in the sample period, manufacturing losses were larger in "manufacturing islands"—counties with strong manufacturing presences that were *not* nested among other factory-heavy counties. Recall from the bar charts in Figure 6 that manufacturing employment generally decamped from cities for less dense areas early in the sample period. The big manufacturing losses among manufacturing islands are consistent with this pattern, as long as the traditional urban manufacturing counties were surrounded

[32]When thinking about the marginal effect of changing the value of an X variable in a spatial model, we need to take spillovers into account. When there is no spatial lag of the Y variable, it is sufficient to take the sum of the coefficients on own X (β) and lagged X (γ). The interpretation of this *total effect* is the effect on one county's Y if it and all first- and second-order neighbors were to raise their value of X by one. Mathematically, one can also interpret the total effect as the sum of spillover and own-effects if *one* county increases its X, and if we took the average of this cumulative effect across all counties. For more information on total effects in spatial models, see LeSage and Pace (2009).

[33]The appropriateness of examining total effects is demonstrated by the similarity between the total effects and OLS coefficients from commuting zone-level regressions, shown in Figure A.7.

low-manufacturing suburbs. Another way to think of this pattern is "concentration through subtraction" in the reallocation of manufacturing jobs. As we have seen, the manufacturing share of employment declined throughout the sample period, even in rural counties. Yet the rural decline was not as severe as in urban counties. Hence, the decline of factory employment in isolated urban manufacturing counties tended to result in manufacturing becoming relatively concentrated away from cities, in pockets of manufacturing-heavy rural counties.

The concentration-through-subtraction pattern can also be seen in the results from the full manufacturing-share regressions shown in Figure 9. The spatial error terms λ_t graphed in the bottom right panel do decline somewhat from the beginning of the sample period until the mid-1990s. However, at that point, the λ_ts decline rapidly, until they stabilize a few years before the sample end. Statistically, the decline in the spatial error terms after 1995 indicates that pockets of manufacturing-heavy counties that had been relatively insulated from manufacturing declines were no longer so insulated. This is exactly what we would expect given the 3-year manufacturing-growth regression results shown in the top left panel of Figure 15. As shown by the dotted line, the positive effect of a strong neighboring manufacturing presence on future factory-job growth disappears in the mid-1990s. In short, the 2000s wave of manufacturing job-losses was an equal opportunity disemployer, adversely affecting both the manufacturing jobs that remained in urban areas as well jobs in less-dense areas, which had been relatively insulated from job losses up to that point.

6.2 Relation to State-Level Data: The Coastal Boom of the 1980s

The finding that factory job losses were larger in urban areas early in the sample helps us understand employment data from US states during the 1980s, when the economies of coastal states boomed relative to those in the nation's interior. This fact was widely noted at the time; the good performance of Massachusetts in particular—the so-called Massachusetts Miracle—was widely believed to propel former governor Michael Dukakis to the 1988 Democratic presidential nomination. The results above imply that factory-job losses in the (relatively urban) Northeast would be especially large, even during the "miracle" years of the 1980s. The blue and red lines in the top panel of Figure 16 plots the manufacturing-employment shares starting in 1960 for the two Census divisions included in the Northeast Census region: New England and the Middle Atlantic states.[34] The heavy black line depicts the manufacturing share for the nation as a whole. In 1960, both divisions have manufacturing shares that are higher than the overall US share. But both these shares decline in the first few decades of the sample, with a particularly significant step-down in the 1980s.

[34]The New England division includes Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. The Middle Atlantic division is comprised by New Jersey, New York, and Pennsylvania.

The lower panel of Figure 16 plots the same data for the other Census divisions east of the Mississippi: East North Central (which includes Michigan, Ohio, and other industrial states of the Upper Midwest), South Atlantic (which includes most of the Eastern Seaboard below New Jersey), and East South Central (which includes the Deep Southern states of Alabama, Kentucky, Mississippi, and Tennessee). For the most part, manufacturing shares in these three divisions declined throughout the sample period at the same pace as the national share. The lone exception is the deep South, where the manufacturing share actually rises until about 1970.

The strong economic performance of the coastal states during the 1980s would be expected to reduce manufacturing's share of state-level employment, simply because job gains outside of the factory sector were so strong. Some data from four selected states shows that there is more to the story. Figure 17 shows that in Massachusetts and New York, manufacturing employment was falling *in absolute terms* during the 1980s boom. By contrast, in two selected states outside the Northeast, North Carolina and Alabama, manufacturing employment was either stable or trended upward for most of the 1974–1999 period. The lower panel of the figure depicts total and manufacturing employment for the same four states from 2000 to 2018. Absolute manufacturing losses in period are far more uniform, consistent with the widespread losses across both urban and rural areas noted above.

6.3 Discussion: Diversification and Human Capital

So why did the national manufacturing decline seem to have larger local effects after 2000? The answer to this question involves two related qualities of a local area: its level of industrial diversification and its level of human capital. Regarding diversification, our results indicate that in the first part of the sample period, manufacturing was becoming increasingly concentrated in contiguous pockets of counties outside of major cities. In particular:

- The bar charts of Figure 6 display the rapid deindustrialization of dense counties throughout the sample, with less of a decline in factory employment in less-dense areas over the same decades;

- The spatial error terms of the manufacturing-share regressions, depicted in Figures 8 and 9, indicate a very high degree of positive spatial correlation among county-level manufacturing-employment shares until about 1995; and

- The manufacturing-employment growth regressions shown in Figure 15 indicate that counties that were spatially isolated in terms of their manufacturing presence experienced much greater factory-job losses than counties embedded among other high-manufacturing counties.

Each of these complementary findings indicates that by the start of our sample period, manufacturing had begun to leave the dense centers of cities. As noted by Wilson (1987, 1996) and others, manufacturing losses in centers of urban areas often had disastrous effects on the workers who lost those manufacturing jobs. The urban areas themselves also suffered in ways that are sometimes hard to remember. Today, New York is often held up as a superstar city while Detroit is closer to the other end of the spectrum. But Glaeser (2011) writes the exodus of automobile jobs from Detroit and garment-industry jobs from New York left both cities in similar straits:

> As recently as the 1970s, pretty much every older industrial city seemed simultaneously doomed. Both New York and Detroit were reeling from the decline of their core industries, and if anything, New York seemed worse off because the car industry seemed more tightly tied to Motown than the garment sector did to Gotham. In 1977, workers in Wayne County, Michigan, which includes Detroit, were paid more than workers in Manhattan. New York City's government didn't seem any better than Detroit's. In 1975, New York State established the Municipal Assistance Corporation to take over the city's finances and stop it from falling into bankruptcy, despite have some of the nation's highest taxes (p. 56).

Some dense cities, like New York, were able to transform there employment bases successfully after losing manufacturing jobs in the 1970s and 1980s. Unfortunately, the spatial pattern of job losses changed in the 1990s in ways that make future success stories like New York less likely. There were big losses both among dense counties *and* among outlying manufacturing clusters that had less potential for diversification. This change is illustrated by the similar percentage declines in manufacturing jobs after 2000 experienced by all four county-density groups in the bar charts of Figure 6; by the post-1995 decline in the spatial error terms from the manufacturing-share regressions presented in Figures 8 and 9; and by the drop in the positive effect of nearby factory employment on manufacturing employment-growth after 1995, as seen in Figure 15. All of these findings are consistent with a pattern in which pockets of outlying, less-dense counties with significant manufacturing shares are hit hard by manufacturing declines.

Unlike larger cities, however, the outlying manufacturing clusters probably had less chance of diversifying away from manufacturing, so the local effects of manufacturing decline on nonmanufacturing job growth were likely to be larger. These rural manufacturing clusters were more vulnerable due to their size; a shock that devastates a rural economy would go unnoticed in a large city. Being more specialized, they also had fewer alternate sources of labor demand to absorb displaced workers. Human capital in particular is a key factor in the resiliency of cities to manufacturing shocks. Economists have long known that education, creativity, and the exchange of ideas lie at the heart of what makes a city successful. The seminal paper of Glaeser et al. (1992), for example, showed that well-diversified cities grew more between 1956 to 1987. The implication is that cities were able to foster the growth through the cross-fertilization of ideas across industries and well as within industries.[35] Human capital is also central to the research outlined in Moretti (2012), which describes a Great Divergence between entrepreneurial "brain hubs" such as San Francisco and Boston and cities with lower education levels.

How do our results relate to the importance of human capital? As pointed out by Eriksson et al. (2019), US manufacturing in the late 20th century may have provided a clear example of the product-cycle model, which is based on differences in human capital across places. In the original framework of Vernon (1966), new products are developed in advanced, innovative countries with high education levels, but over time, production becomes standardized and migrates to countries with less human capital and lower wages. An insight of Eriksson et al. (2019) is that Vernon's model can be applied to localities within the United States as well as across counties. Specifically, the migration of manufacturing employment into less-educated areas during the last half of the 20th century provides a close parallel to the growth of manufacturing in low-wage countries. The product cycle could help explain why the effects of the China shock analyzed by Autor, Dorn, and Hanson (2013) turned out to be so much larger than economists would have expected. After China joined the World Trade Organization, low-cost Chinese imports began to compete directly with low-wage US regions that were at a similar stage of the product cycle. Previous trade shocks, on the other hand, had emanated from Japan and the so-called Asian Tigers, where wages were somewhat higher. Eriksson et al. (2019) contend that because the communities affected by these previous shocks were earlier in the product cycle, and generally better educated, they could withstand those shocks more easily.

A number of our empirical findings buttress the product-cycle argument; that is, we find evidence not only that manufacturing rotated from dense to less-dense areas, but also that these areas tended to have lower education levels and wages. Perhaps the best evidence comes from a comparison of the manufacturing-share regressions in Figures 8 and 9. The first of these figures depicts coefficients from the regressions that include density-related terms but omit the share of residents with bachelor's degrees. When education is included in the regression in Figure 9, the importance of the density term is reduced. The implication is that much of the migration of factory employment to less-dense areas is more accurately described as a migration to less-educated areas. Additionally, the path of yearly coefficients on the college-attainment term in Figure 9 suggests that although less-educated areas were generally more successful in attracting or retaining manufacturing employment early in the sample period, they were less so after 1995—close to the time of the China shock.

[35] See also Glaeser (2011) on this point.

Other evidence for a rotation towards low-wage areas comes from the average earnings regressions in Figure 13. As noted, the share-manufacturing term is always positive, consistent with the high wages generally paid to manufacturing workers. But the neighboring-share term is negative, indicating that manufacturing plants tend to be located in low-wage areas. This term becomes more negative over the first part of the sample period, indicating higher growth of manufacturing in lower-wage parts of the country. Interestingly, the mid-1980s turnaround in these coefficients occurs about ten years earlier that the reversal of the education coefficients in the average-earnings regression.[36] Taken together, however, our empirical results suggest that the communities hit by recent manufacturing declines were less able to diversify because they were not particularly dense and because they had lower education levels.

7 Some Evidence Regarding Within-City Inequality

Policies to deal with geographic disparities follow two broad outlines. Some are designed to facilitate the migration of people from declining areas into successful ones, often by reducing barriers to residential construction in thriving cities.[37] Others, known as place-based policies, encourage growth in less-favored localities. Although economists have traditionally been wary of such policies, new ideas about helping lagging cities are becoming an important part of the policy debate. For example, Gruber and Johnson (2019) argue that aggressive government support for basic research and development in several mid-sized American cities would not only boost overall US productivity growth, but also rebalance economic opportunity across different locations. And Austin, Glaeser, and Summers (2018) contend that direct monetary subsidies for job creation could be socially worthwhile in places where the employment rates are currently low.

An empirical fact that bears on this policy choice is that earnings inequality *within* large and dense cities has recently increased. Baum-Snow, Freedman, and Pavan (2018) use a structural model to study this phenomenon, and the recent Ely Lecture by Autor (2019) links it to fundamental changes in the nature of job opportunities available in large cities.

[36] For some firm-level evidence on the recent behavior of manufacturing employment, see Bloom et al. (2019) and Fort, Pierce, and Schott (2018).

[37] Ganong and Shoag (2017) link the cessation of regional income convergence in the United States to restrictive land-use regulations that housing construction in high-income areas. Hsieh and Moretti (2019) note that such restrictions are particularly severe in three metropolitan areas with strong recent productivity growth: New York, San Francisco, and San Jose. The authors contend that reducing constraints in those cities alone would raise the US GDP growth rate by more than one third. Glaeser and Gyourko (2018) show that although estimates of the social cost of housing restrictions are typically sensitive to model assumptions, "[t]he available evidence suggests, but does not definitively prove, that the implicit tax on development created by housing regulations is higher in many areas than any reasonable negative externalities associated with new construction" (p. 5).

Autor presents data illustrating the changes over time in the relationship between average wages and local population density for different types of workers.[38] One set of exhibits splits workers on the basis of education, specifically, workers with high-school educations or below and those with some years of college and above. In 1970, wages for both groups rise sharply with city-level density. But by 2015, the end of Autor's sample period, the density premium for non-college workers non-college workers had nearly vanished. In other words, today's college workers tend to earn more if they work in dense cities, as they always have. But less-educated workers no longer do, and earnings inequality within large cities is higher as a result.

This finding informs policy choices because it suggests (to use Autor's wording) that dense cities are no longer the "land of opportunity" for lower-skilled workers. If so, then the reduced migration of such workers to big cities may not result from higher costs of housing there, but rather a lack of job opportunities, so that policies intended to encourage migration may not have their desired effects. In this subsection, we outline how theories of the labor market have been applied to the issue of rising urban inequality, and how the empirical work used to confront those theories relates to our findings above. As we will see, fundamental changes in the organization of work since the start of the 21st century—particularly the tasks done by secretaries, clerks, and other types of office-support personnel—will play an important part in explaining the inequality finding.

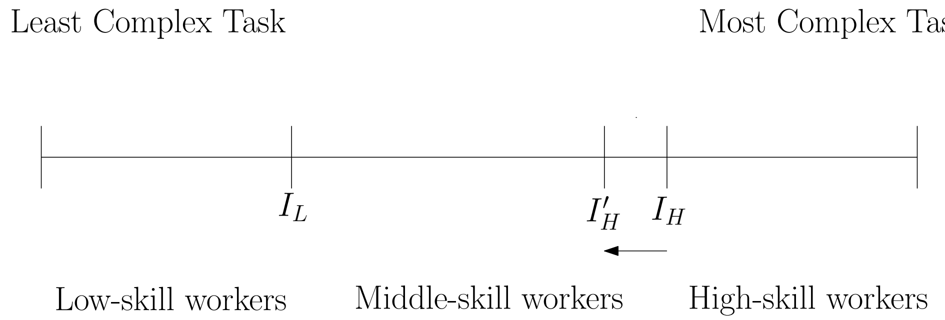
Autor (2019) links the inequality result to recent theoretical work on wage determination, which has flourished as concern about earnings inequality is increased. What is sometimes called the canonical model of the labor market starts by classifying workers on the basis of skill, often defined on the basis of college attendance or degree attainment. The labor hours supplied by different groups enter into a single production function that allows for technical change to be biased towards high-skill workers, that is, technical change can raise the productivity of high-skill workers more quickly than the productivity of lower-skill workers.[39] Combined with data on the relative supplies of workers in different skill groups over time, this model does a good job of explaining the rapid increase in the college premium at the start of the 1980s (Goldin and Katz 2008).

Yet the canonical model does less well in accounting for other dimensions of the wage distribution, most notably the decline of real wages *in absolute terms* for some less-educated groups. To expand the explanatory power of the model, Acemoglu and Autor (2011) expand it to include a layer of "tasks" between labor supply and output production. Whereas the canonical model directly maps the labor supplied by different groups of workers to a certain

[38] See in particular Figure 13 in Autor (2019). The spatial unit of observation in throughout Autor's paper is the commuting zone.

[39] A related strategy is to posit a complementary between capital and skill that also boosts the productivity of high-skill workers in relative terms.

level of output, the expanded model maps workers in different skill groups to different sets of tasks, and it is the performance of these tasks that creates output.



The Direct Effect of a Skill-Biased Change in Task Assignment in Acemoglu and Autor (2011).

The above schematic presents the nucleus of what Acemoglu and Autor (2011) call a "Ricardian" model of the labor market. The model includes three skill groups, and workers in these groups perform tasks of varying complexity. The line in the figure represents the distribution of potential tasks, with the most complex task appearing on the right end of the line and the least complex on the left. The authors show that under plausible assumptions, workers in the three skill groups will specialize in the performace of tasks for which they have a comparative advantage (hence the Ricardian nature of the model). Because high-skill workers can perform complex tasks more efficiently than workers in the other two groups, high-skill workers specialize in the tasks that lie to the right of I_H, the border that divides their tasks from those performed by middle-skill workers. The middle-skill workers perform the intermediate tasks between I_H and I_L, with the low-skill workers performing the tasks to the left of I_L.

The schematic also depicts a leftward movement in I_H to I'_H, indicating that the set of tasks performed by high-skill workers has grown. One reason this can occur is technical change. For example, a new invention in the hands of high-skill workers might allow them to perform tasks just to the left of the previous I_H more efficiently than the middle-skill workers can. Even though high-skill workers are always paid more than middle-skill workers (another implication of the model), the new invention could boost the productivity of the high-skill workers in tasks just to the left of I_H by enough to make high-skill workers the most efficient group to perform those tasks. If so, then the set of high-skill tasks will expand as I_H moves leftward, leaving the group of middle-skill workers to be squeezed onto a smaller set of task opportunities. The resulting reduction of middle-skill labor demand can reduce their real wages in absolute terms.[40]

[40] Because the reduced wages of the middle-skill group makes them cheaper to employ for all tasks, there would be an offsetting leftward movement in I_L (not shown in the figure) as the middle-skill workers take

In empirical work, two broad classes of middle-skill workers are production workers (who are often manufacturing workers) and office and administrative-support workers (who work in a variety of industries). Autor (2019) shows that between 1970 and 1990, there was a large decline in the share of big-city employment accounted for by production workers, with a smaller decline in office and administrative-support employment. Between 1990 and 2015, however, there is not only a continued decline in production work in big cities but also a large decline in office and administrative work. As suggested by our schematic drawing, the Ricardian model can explain this pattern as resulting from technical change. Word-processing software, computer spreadsheets, large databases, and other technical innovations since the 1980s may have allowed high-skill workers to absorb the clerical and secretarial tasks that had previously been performed by middle-skill workers. As Autor (2019) explains, the resulting squeeze in middle-skill job opportunities would then force non-college workers to migrate down the task-complexity distribution to perform lower-skill, service-oriented tasks. Because these tasks pay less than the lost office-and-administrative tasks, this migration increases wage inequality in dense urban areas.

Our results are consistent with this line of thinking in several ways. First, as noted above, production workers are disproportionately employed in manufacturing. Thus our documentation of big reductions in employment in the manufacturing *industry* in dense counties early on is consistent with Autor's finding of large losses in the production-worker *occupations* over the same time period. Because the CBP data is industry-based, however, it is not as well-suited to study office-and-administrative employment, an occupation that experienced large losses later in the sample period. Consequently, we draw on work by Foote and Ryan (2014), who used historical data from the Current Population Survey (CPS) to construct consistent occupational shares for different skill groups from the late 1940s onward. The top panel of Figure 18 depicts employment shares for four broad occupational classes, which are defined using the routine/nonroutine and cognitive/manual nature of their typical tasks. Nonroutine cognitive work includes management and professional jobs and is considered to be high-skill work, while routine cognitive occupations include office and administrative work and sales positions and is usually considered middle-skill work. Routine manual work primarily consists of production jobs, often in manufacturing, and is also considered middle-skill, while the nonroutine manual jobs denote lower-skill jobs, primarily in service occupations.

The red line in the panel illustrates the steady increase in the high-skill, nonroutine cognitive employment during the second half of the 20th century, which is a central focus of

some jobs previously done by low-skill workers. But Acemoglu and Autor (2011) show that this leftward movement is only partially offsetting—the main effect of the technical change is to make the middle-skill workers worse off. For other applications of the model to the study of technical change, particularly robotics, see Acemoglu and Autor (2011) and Acemoglu and Restrepo (2018, 2019).

Goldin and Katz's (2008) study of how long-run technical progress has steadily boosted the demand for skilled workers. The green line shows the long decline in the share of routine manual employment, consistent with the relative decline in the manufacturing industry over this period. For our purposes, the most important line is the dark blue one, as it shows a decline in the share of office-and-administrative jobs and other routine cognitive occupations that begins around 1990. The timing of this decline lines up well with the decline in office-and-administrative positions that Autor highlights using Census data.

The 1990 timing for this decline also lines up with the decline in relative non-college *wages* in big cities explored in Autor (2019), as opposed to changes in occupational shares. That is, there were big occupational shifts for non-college workers in big cities before 1990, driven in large part by the migration of blue-collar manufacturing jobs out of dense cities in the 1960s and 1970s. But during this time, enough opportunities for non-college workers seem to have emerged elsewhere to keep the non-college wage-density gradient largely in place—until the 1990s. Some additional data from the Census and ACS shows that many of these jobs were in the office support category. The lower panel of Figure 18 graphs annualized intercensal percentage changes in the growth of clerical and office-support employment starting with the 1950 Census, using the density-group classifications we developed earlier. In the top 1% of counties, growth in office-support employment is indeed lower throughout the sample period. But consistent with the national data in the top panel of the figure, there is substantial growth in office-support employment everywhere else until the 1990s.

8 Conclusion

This paper has taken a broad look at changes in employment and earnings across US counties since 1964, in order to shed light on how geographic disparities in labor markets have emerged and what policies might best address them. Perhaps the best way to sum up our results is to note the parallels and contrasts of the current geographic disparities with similar developments in previous decades.

Starting with the parallels: the early CBP data drive home the fact that the consequences of manufacturing job declines is nothing new. It is easy to forget that the exodus of manufacturing jobs from big cities near the middle of the 20th century caused significant problems both for the affected workers and for the cities themselves. Smaller cities experienced less-severe manufacturing losses during this period, and in some cases substantial absolute increases in manufacturing employment. Our regression analysis indicates that this spatial pattern changed in the 1990s, with outlying manufacturing clusters of counties experiencing large losses as well. The challenges now faced by these more-rural areas closely parallel the challenges faced by New York, Detroit, and other large industrial cities in previous years. Unfortunately, the location of these outlying manufacturing areas, as well as their lower levels of human capital, may make them less able to diversify themselves in response to the manufacturing decline. As noted by Eriksson et al. (2019), the lower rung of the product cycle on which these areas find themselves may explain why the China shock studied in Autor, Dorn, and Hanson (2013) has had such long-lasting effects.

Yet there are substantial contrasts between the current situation and past decades as well, most of which are no doubt linked to ongoing technological change. For example, for most of the postwar era there was a pronounced trend toward decentralization and suburbanization in large metropolitan areas. For areas with the densest core counties, however, the decentralization of both population and employment appears to have slowed significantly or even stopped since the turn of the 21st century. Evidence for cities with somewhat less-dense cores is more ambiguous. Do these patterns indicate that thanks to technological growth, agglomeration effects for high-skill workers in cities have increased (Baum-Snow, Freedman, and Pavan 2018)? Or are these patterns driven by the consumer amenities in cities instead (Couture et al. 2019)?

A related finding is that the correlation of county-level density with average earnings has increased, as a relationship resembling a hockey stick has emerged between density percentiles and average earnings. In addition, the dispersion in average earnings at a given density percentile, which had been falling or stable for much of the past several decades, has increased markedly since the late 1990s for the nation's densest counties. These earnings patterns at the county level are no doubt related to earnings developments at the individual level. Indeed, new models designed to explain the distribution of income for the nation as a whole suggest that job opportunities for less-educated workers in big cities have dried up. The critical question for policymakers is whether these trends now justify more aggressive policies to spread employment opportunities more widely across the country.

References

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, Orley Ashenfelter and David E. Card, eds., vol. 4, Part B, 1043–1171. Amsterdam: North Holland/Elsevier.

Acemoglu, Daron, and Pascual Restrepo. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review* 108(6): 1488–1542.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives* 33(2): 3–30.

Austin, Benjamin A., Edward L. Glaeser, and Lawrence H. Summers. 2018. "Jobs for the Heartland: Place-Based Policies in 21st Century America." *Brookings Papers on Economic Activity* 2018(1): 151–232.

Autor, David H. 2019. "Work of the Past, Work of the Future." *AEA Papers and Proceedings* 109: 1–32.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103(6): 2121–2168.

Baum-Snow, Nathaniel, Matthew Freedman, and Ronni Pavan. 2018. "Why Has Urban Inequality Increased?" *American Economic Journal: Applied Economics* 10(4): 1–42.

Bloom, Nicholas, Kyle Handley, Andre Kurman, and Philip Luck. 2019. "The Impact of Chinese Trade on U.S. Employment: The Good, The Bad, and The Debatable." Working Paper. Available at https://nbloom.people.stanford.edu/sites/g/files/sbiybj4746/f/bhkl_posted_draft.pdf.

Charles, Kerwin Kofi, Erik Hurst, and Mariel Schwartz. 2018. "The Transformation of Manufacturing and the Decline in US Employment." In *NBER Macroeconomics Annual 2018*, Martin Eichenbaum and Jonathan Parker, eds., vol. 33, 307–372. Chicago: University of Chicago Press.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst. 2019. "Income Growth and the Distributional Effects of Urban Spatial Sorting." National Bureau of Economic Research Working Paper No. 26142.

Eriksson, Katherine, Katheryn Russ, Jay C. Shambaugh, and Minfei Xu. 2019. "Trade Shocks and the Shifting Landscape of US Manufacturing." National Bureau of Economic Research Working Paper No. 25646.

Florida, Richard. 2017. *The New Urban Crisis: How Our Cities Are Increasing Inequality, Deepening Segregation, and Failing the Middle Class—and What We Can Do About It.* New York: Basic Books.

Foote, Christopher L., and Richard W. Ryan. 2014. "Labor-Market Polarization over the Business Cycle." In *NBER Macroeconomics Annual 2014*, Jonathan Parker and Michael Woodford, eds., vol. 29, 371–413. Chicago: University of Chicago Press.

Fort, Teresa C., Justin R. Pierce, and Peter K. Schott. 2018. "New Perspectives on the Decline of US Manufacturing Employment." *Journal of Economic Perspectives* 32(2): 47–72.

Ganong, Peter, and Daniel Shoag. 2017. "Why Has Regional Income Convergence in the US Declined?" *Journal of Urban Economics* 102: 76–90.

Glaeser, Edward L. 2011. *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier.* New York: The Penguin Press.

Glaeser, Edward L., and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives* 32(1): 3–30.

Glaeser, Edward L., and Matthew E. Kahn. 2001. "Decentralized Employment and the Transformation of the American City." *Brookings-Wharton Papers on Urban Affairs* 2001: 1–63.

Glaeser, Edward L., and Matthew E. Kahn. 2004. "Sprawl and Urban Growth." In *Handbook of Regional and Urban Economics*, J. Vernon. Henderson and Jacques-François Thisse, eds., vol. 4, 2481–2527. Amsterdam: Elsevier.

Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100(6): 1126–1152.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. "Consumer City." *Journal of Economic Geography* 1(1): 27–50.

Goldin, Claudia D., and Lawrence F. Katz. 2008. *The Race Between Education and Technology.* Cambridge, MA: Belknap Press of Harvard University Press.

Gruber, Jonathan, and Simon Johnson. 2019. *Jump-Starting America: How Breakthrough Science Can Revive Economic Growth and the American Dream*. New York: PublicAffairs Books.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst. 2013. "Endogenous Gentrification and Housing Price Dynamics." *Journal of Public Economics* 100: 45–60.

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11(2): 1–39.

LeSage, James, and Robert Kelley Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Chapman and Hall/CRC.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. 2019. "IPUMS National Historical Geographic Information System: Version 14.0 [Database]." Available at http://doi.org/10.18128/D050.V14.0.

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

Ody, Christopher John, and Thomas N. Hubbard. 2011. "County Business Patterns, 1962, 1964-1970: U.S. Summary, State, and County Data." Available at https://doi.org/10.3886/ICPSR25984.v2.

Vernon, Raymond. 1966. "International Investment and International Trade in the Product Cycle." *Quarterly Journal of Economics* 80(2): 190–207.

Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. First edition. Chicago: University of Chicago Press. Second edition published in 2012.

Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Alfred A. Knopf.

Winkler, Richelle, Kenneth M. Johnson, Cheng Cheng, Jim Beaudoin, Paul R. Voss, and Katherine J. Curtis. 2013. "Age–Specific Net Migration Estimates for US Counties, 1950–2010." Applied Population Laboratory, University of Wisconsin, Madison, WI. Available at https://netmigration.wisc.edu/.

Density Group Number	Density Percentiles Included	No. of Counties	Share of Pop.	Share of Counties with Denser Neighbor
1	1–85	2,473	$\approx$ 30%	$\approx$ 35%
2	86–95	291	$\approx$ 25%	$\approx$ 50%
3	96–99	116	$\approx$ 30%	$\approx$ 30%
4	100	29	$\approx$ 15%	0%
Total		2,909	100%	

Table 1. Density Groups for Counties.
Source: US Census Bureau (land area, 1960 population); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).

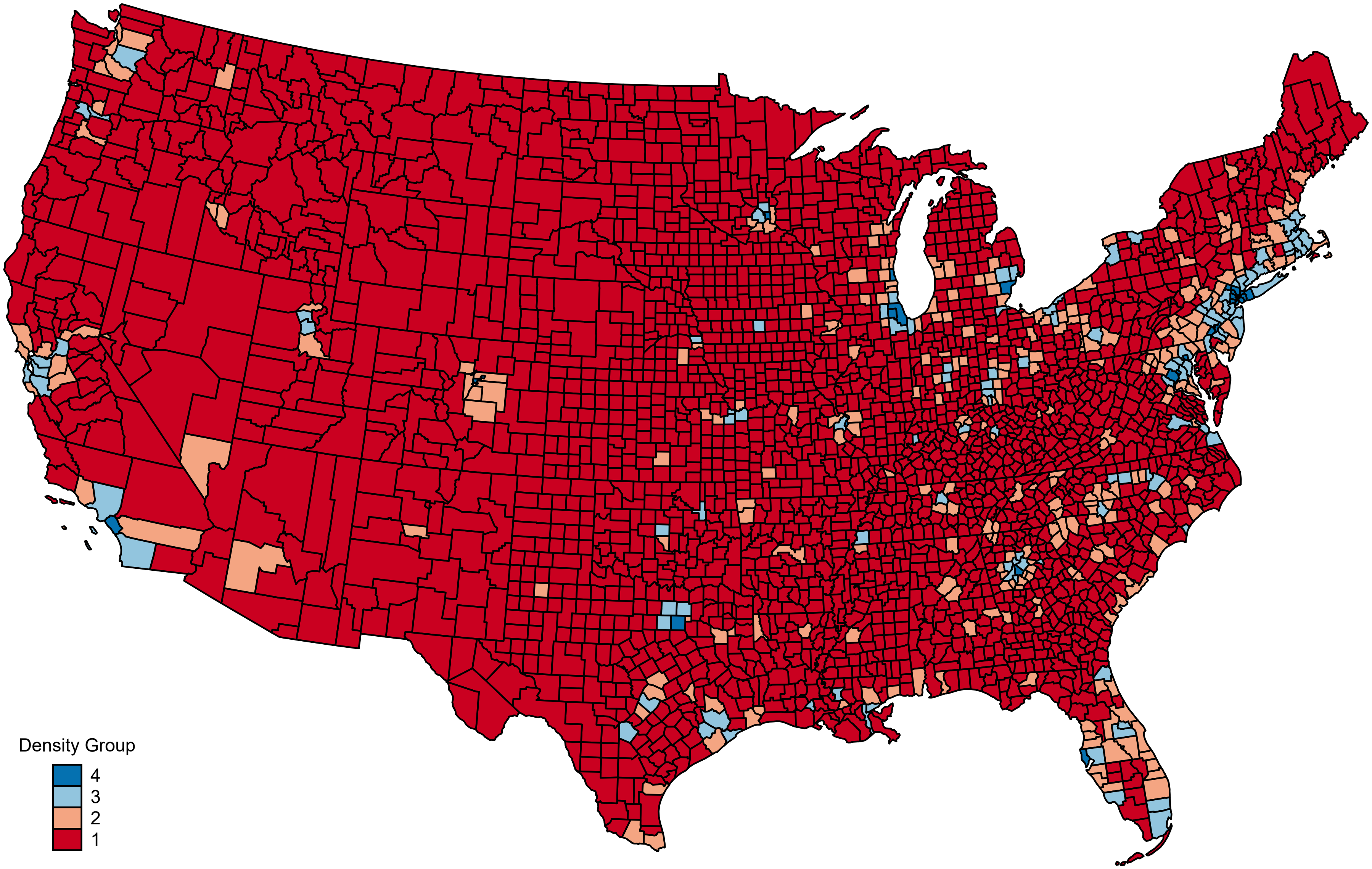


Figure 1. US Counties in 2016, by Density Group.
Note: Counties in Density Group 1 (red) are in percentiles 1–85 of the population density distribution. Those in Group 2 (tan) are in percentiles 86-95, Group 3 (light blue) consists of percentiles 96–99, and Group 4 (dark blue) corresponds to the top percentile of density. Some counties and independent cities have combined; see the text for details. *Source:* US Census Bureau (shapefiles, land area); Surveillance, Epidemiology, and End Results (SEER) program (2016 population);

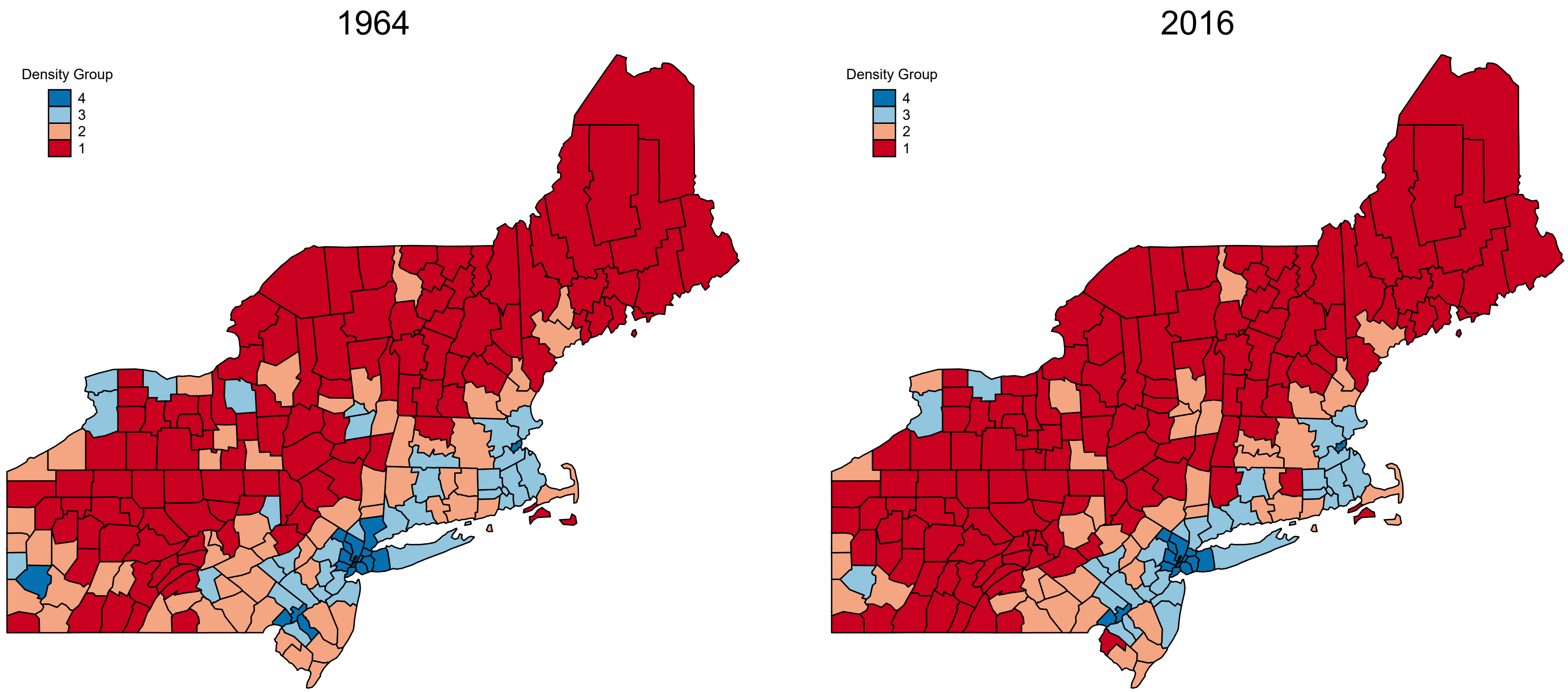


Figure 2. Counties in the Northeast Region in 1964 and 2016, by Density Group. *Note:* See the notes to Figure 1 for details on the classification of counties *Source:* US Census Bureau (land area, 1960 population); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).

Name of County	State	1964	1973	1989	2000	2016
Orange	CA			x	x	x
San Francisco	CA	x	x	x	x	x
Denver	CO	x	x	x	x	x
District of Columbia	DC	x	x	x	x	x
Pinellas	FL		x	x	x	x
DeKalb	GA					x
Cook	IL	x	x	x	x	x
DuPage	IL			x	x	x
Orleans	LA	x	x	x	x	
Baltimore	MD	x	x	x	x	x
Suffolk	MA	x	x	x	x	x
Wayne	MI	x	x	x	x	x
Ramsey	MN	x	x	x	x	x
St. Louis	MO	x	x	x	x	x
Bergen	NJ	x	x	x	x	x
Camden	NJ	x	x			
Essex	NJ	x	x	x	x	x
Hudson	NJ	x	x	x	x	x
Passaic	NJ	x	x	x	x	x
Union	NJ	x	x	x	x	x
Bronx	NY	x	x	x	x	x
Kings	NY	x	x	x	x	x
Nassau	NY	x	x	x	x	x
New York	NY	x	x	x	x	x
Queens	NY	x	x	x	x	x
Richmond	NY	x	x	x	x	x
Westchester	NY	x				
Cuyahoga	OH	x	x	x	x	
Hamilton	OH	x	x			
Allegheny	PA	x	x			
Delaware	PA	x	x	x	x	x
Philadelphia	PA	x	x	x	x	x
Dallas	TX					x
Fairfax - Arlington - Falls Church City Alexandria City - Fairfax City	VA			x	x	x
Milwaukee	WI	x	x	x	x	x

Table 2. Densest 1% of Counties in Selected Years.
Note: Both population and land area are defined by the Census. *Source:* US Census Bureau (land area, 1960 population); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).

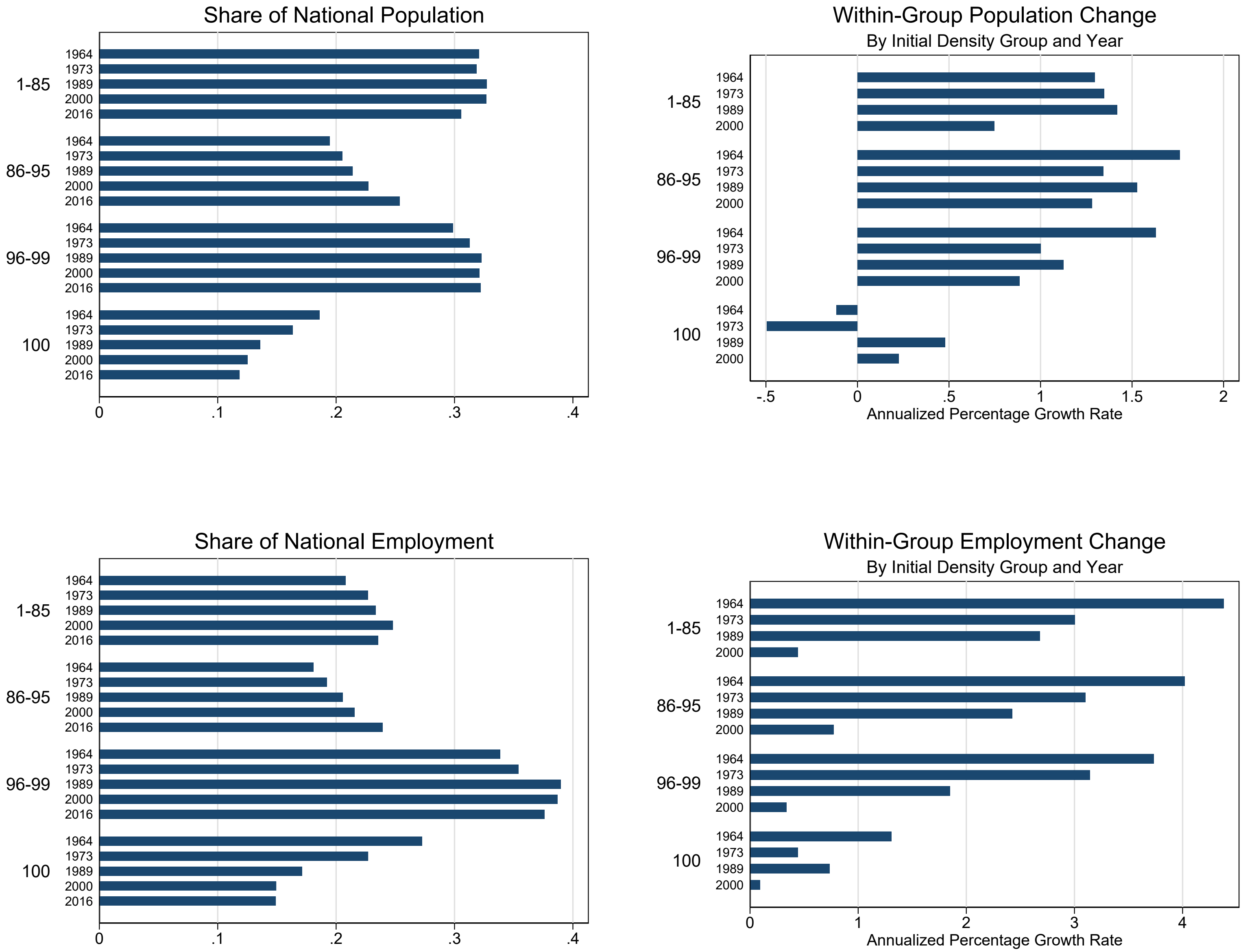
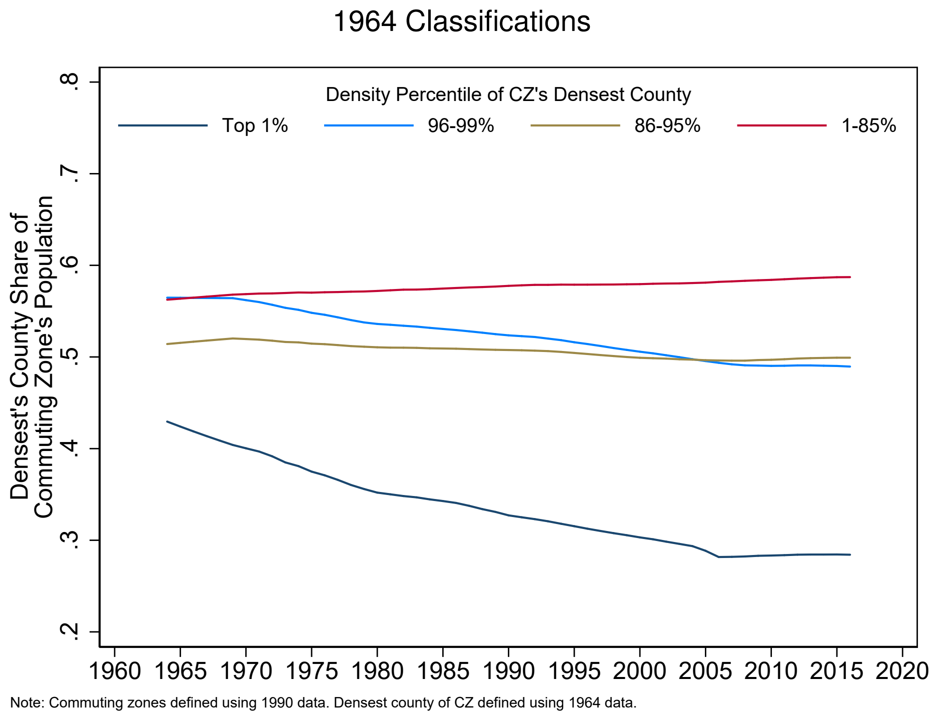


Figure 3. Population, Employment, and Density.
Source: US Census Bureau (land area, 1960 population, 1964-2016 employment); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).

Largest City in 1990 Commuting Zone	Commuting Zone Code	1964	1973	1989	2000	2016
Atlanta, GA	9100					x
Baltimore, MD	11302	x	x	x	x	x
Boston, MA	20500	x	x	x	x	x
Chicago, IL	24300	x	x	x	x	x
Cincinnati, OH	12701	x	x			
Cleveland, OH	15200	x	x	x	x	
Dallas, TX	33100					x
Denver, CO	28900	x	x	x	x	x
Detroit, MI	11600	x	x	x	x	x
Dist. of Columbia	11304	x	x	x	x	x
Los Angeles, CA	38300			x	x	x
Milwaukee, WI	24100	x	x	x	x	x
Minneapolis, MN	21501	x	x	x	x	x
New Orleans, LA	3300	x	x	x	x	
New York, NY	19400	x	x	x	x	x
Newark, NJ	19600	x	x	x	x	x
Philadelphia, PA	19700	x	x	x	x	x
Pittsburgh, PA	16300	x	x			
San Francisco, CA	37800	x	x	x	x	x
St. Louis, MO	24701	x	x	x	x	x
Tampa, FL	6700		x	x	x	x
Total Included CZs		17	18	17	17	17

Table 3. 1990 Commuting Zones with Central Counties among the Densest 1% of Counties. *Note:* Both population and land area are defined by the Census. *Source:* US Census Bureau (land area, 1960 population); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (1990 commuting zone definitions).



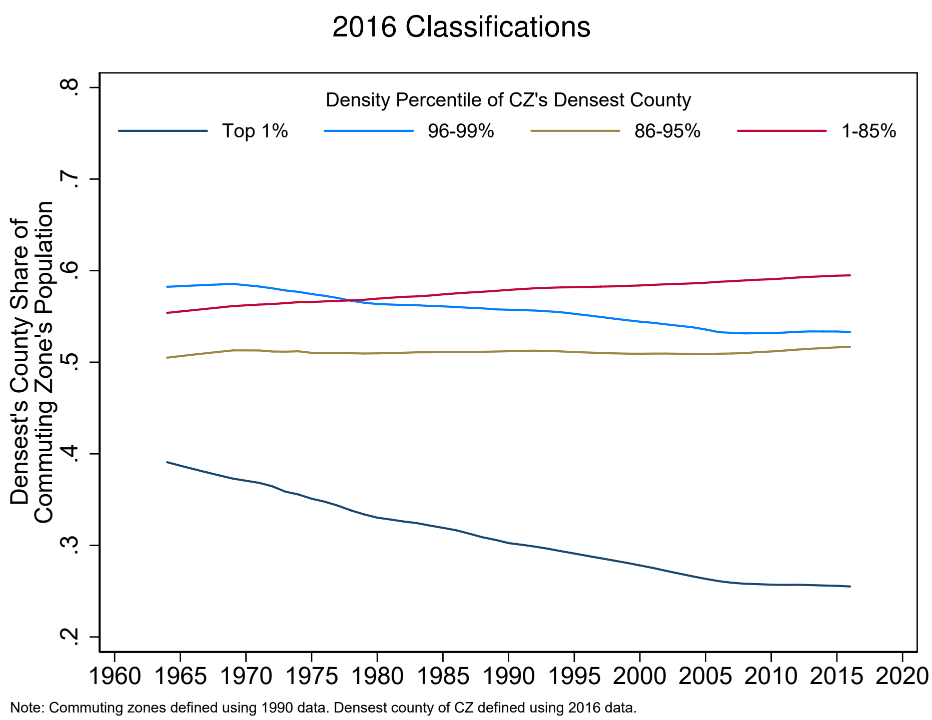
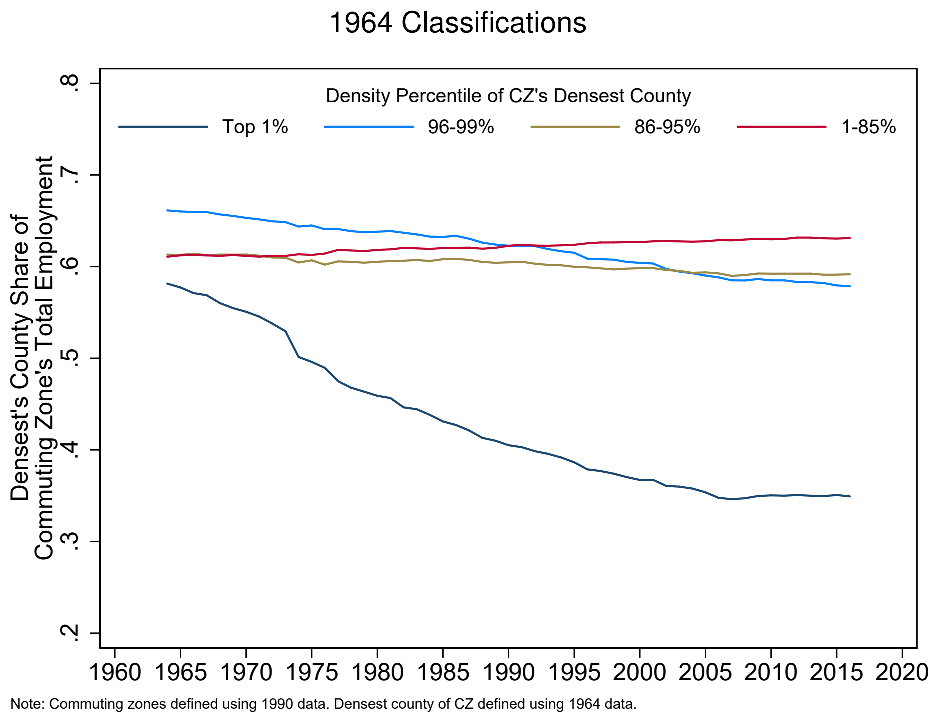


Figure 4. Centralization of Population within Commuting Zones: 1964–2016.
Source: US Census Bureau (land area, 1960 population); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (1990 commuting zone definitions).



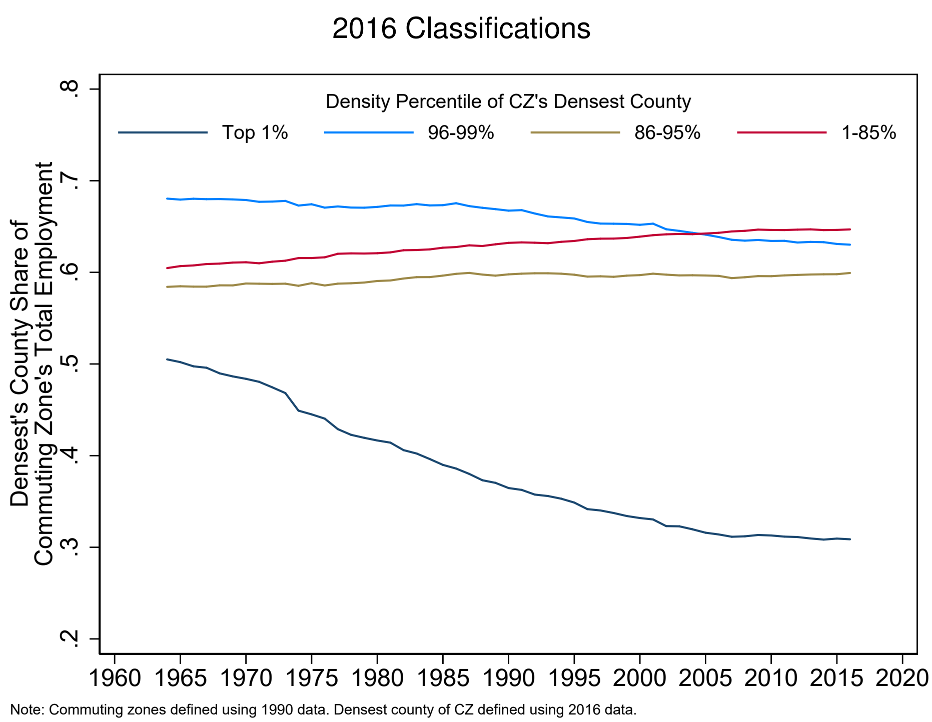


Figure 5. Centralization of Employment within Commuting Zones: 1964–2016.
Source: US Census Bureau (land area, 1960 population, 1964-2016 employment); Surveillance, Epidemiology, and End Results (SEER) program (2016 population); US Department of Agriculture (1990 commuting zone definitions).

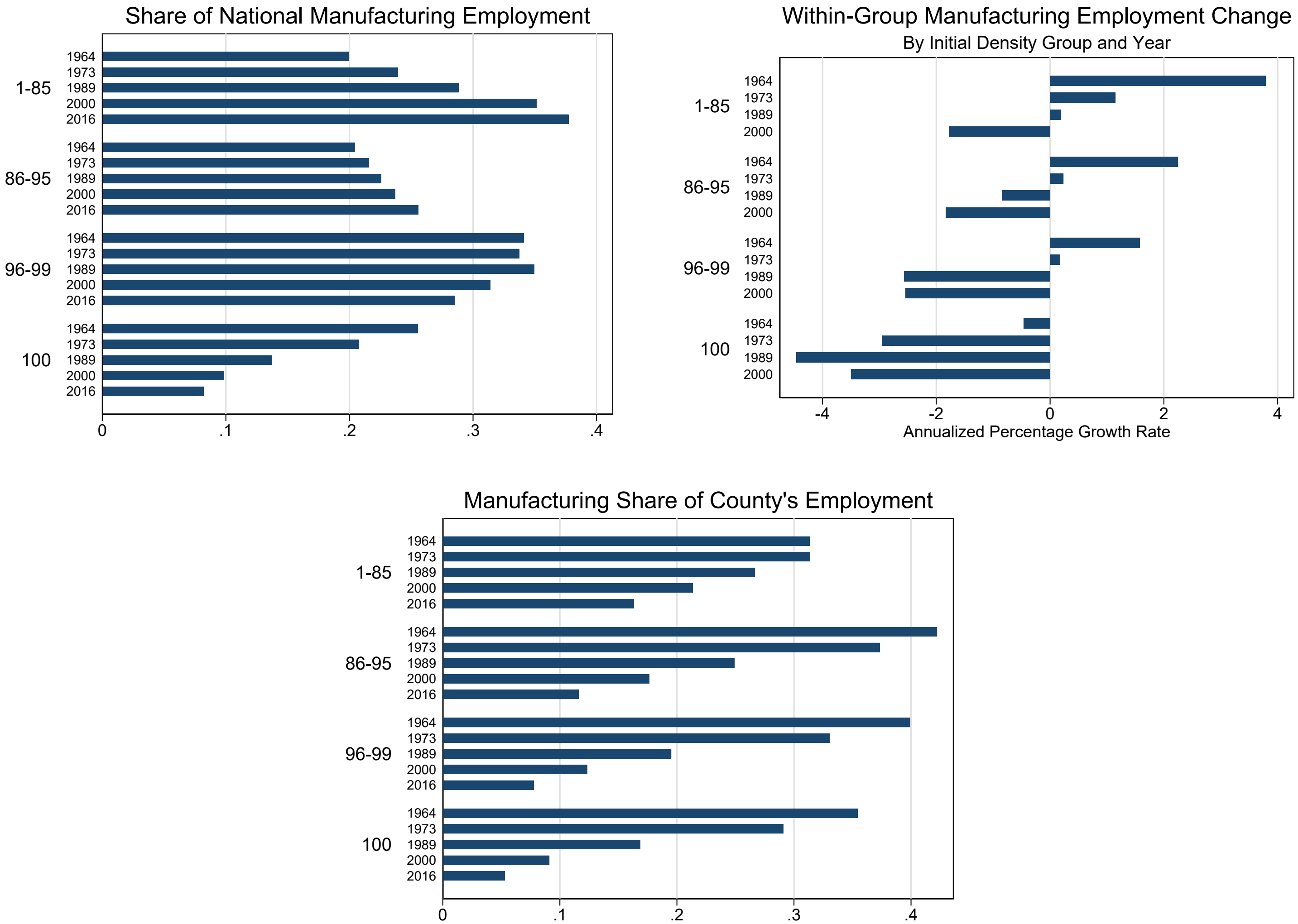


Figure 6. Manufacturing Employment and Density.
Source: US Census Bureau (land area, 1960 population, 1964-2016 total and manufacturing employment); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).

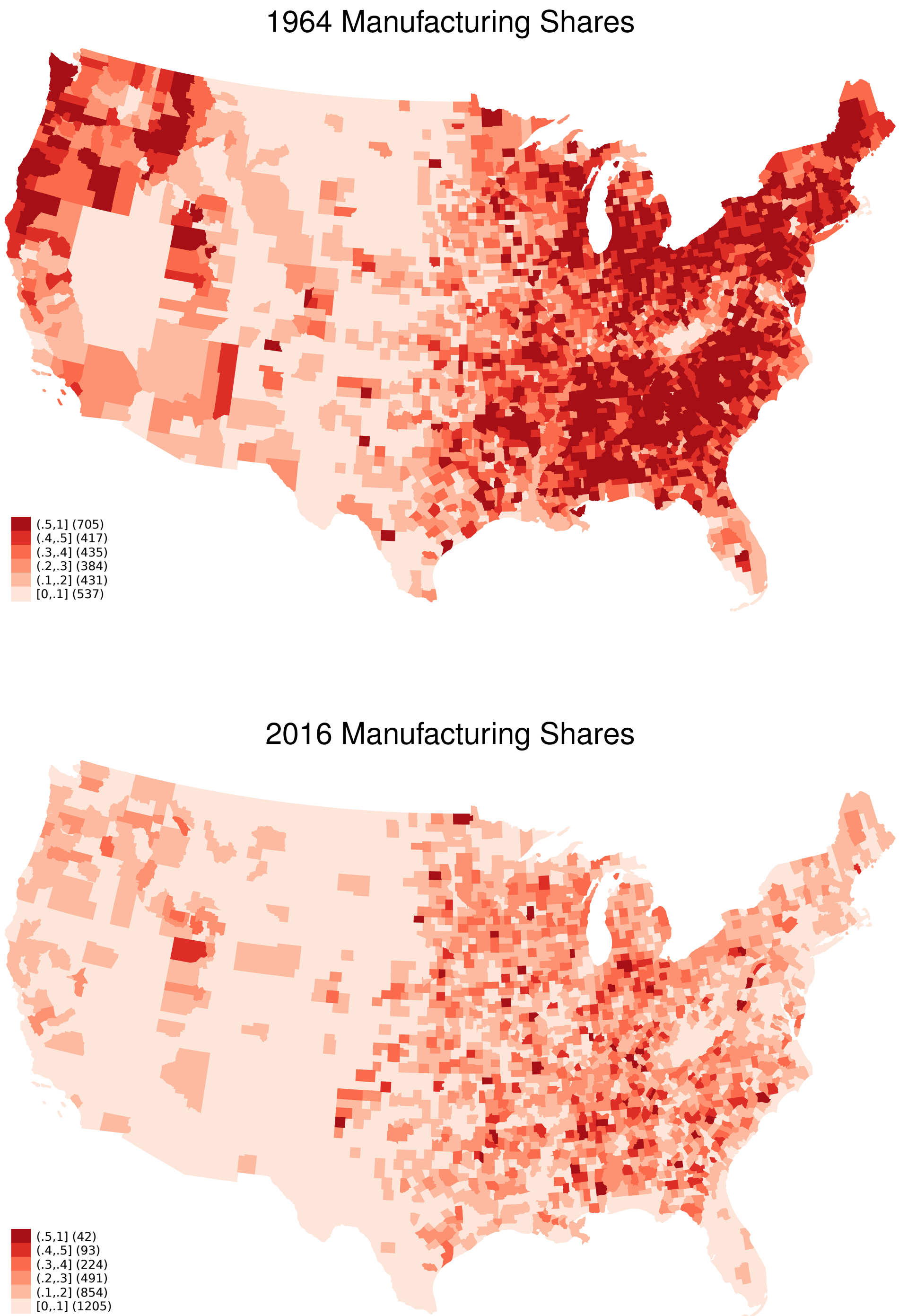


Figure 7. Manufacturing Shares in 1964 and 2016. *Source:* US Census Bureau (shapefiles, 1964-2016 total and manufacturing employment).

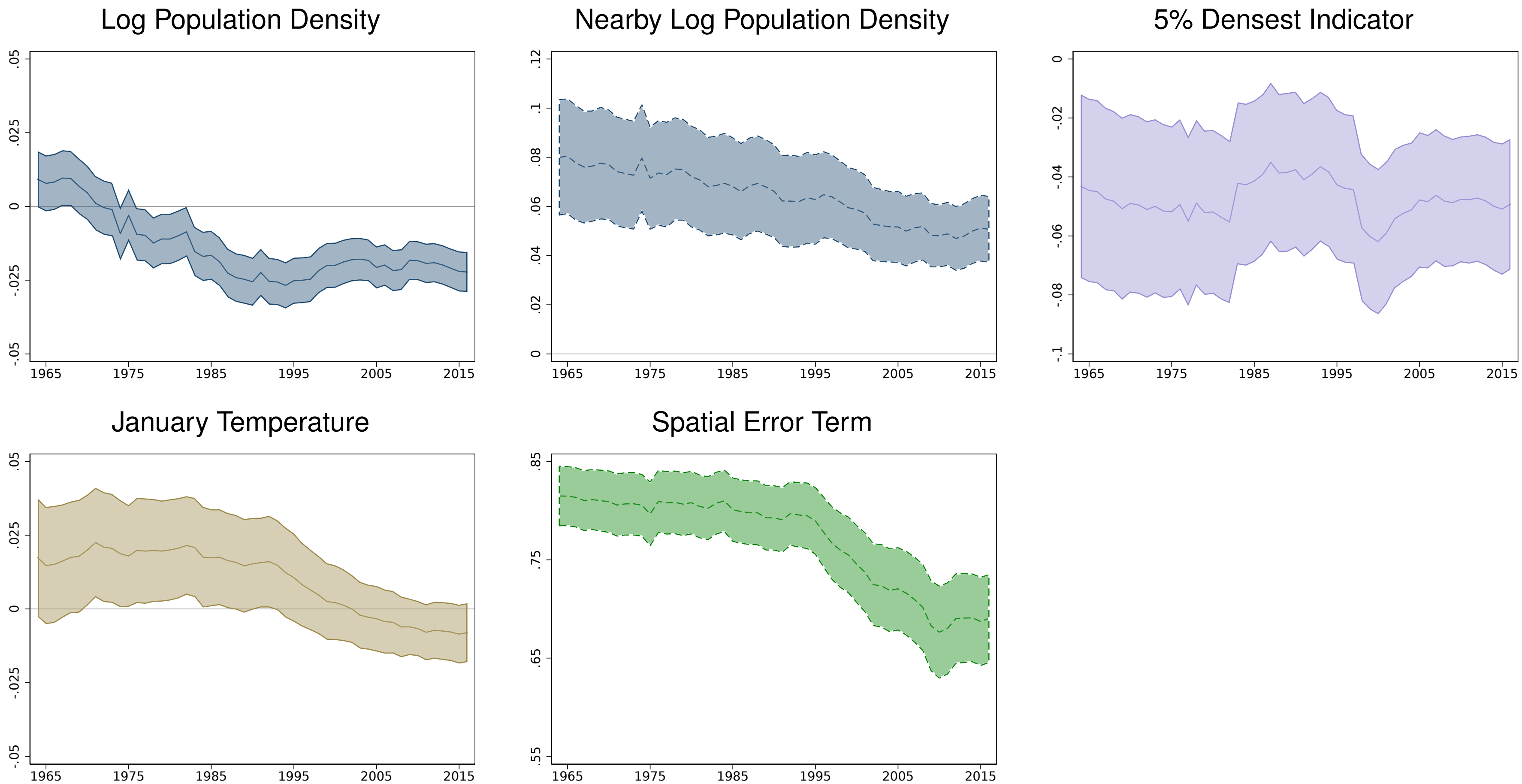


Figure 8. Regression Model for County-Level Manufacturing Shares, Excluding Bachelor's Degrees.
Source: US Census Bureau (land area, 1960 population, shapefiles, total and manufacturing employment 1964-2016); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).

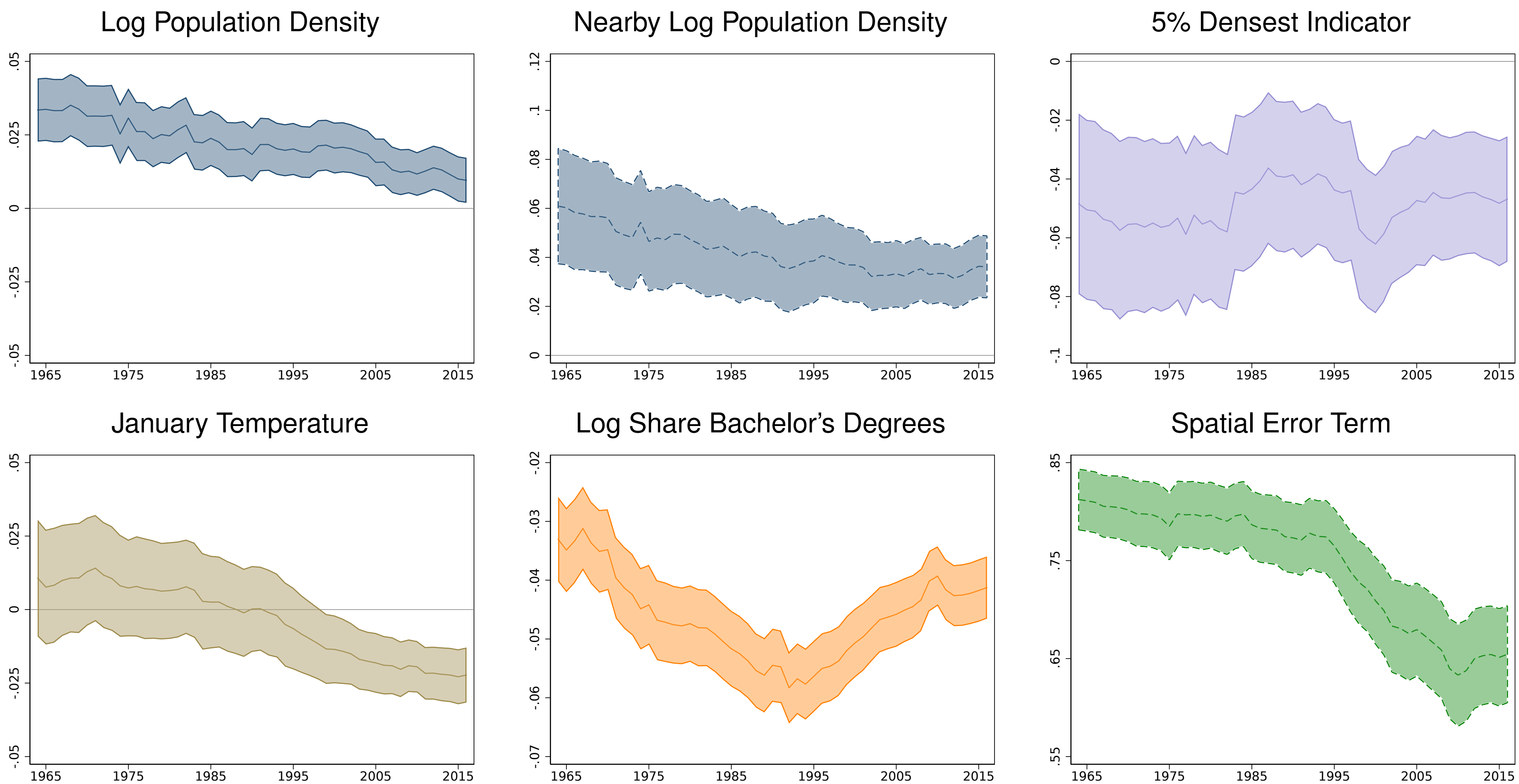


Figure 9. Regression Model for County-Level Manufacturing Shares, Including Bachelor's Degrees.
Source: US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), total and manufacturing employment 1964-2016); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).

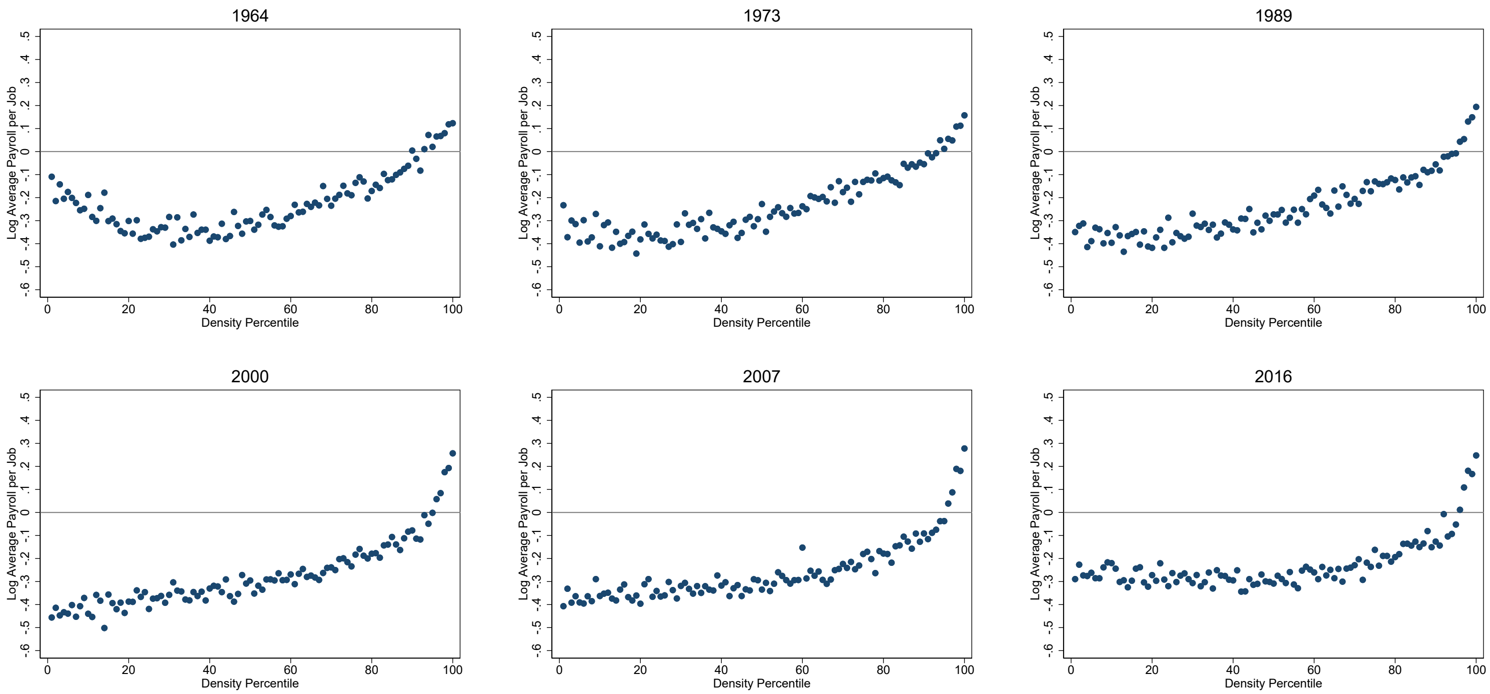


Figure 10. Average Earnings per Job Relative to Population-Weighted County Mean, by Density Percentile and Year. *Source:* US Census Bureau (land area, 1964-2016 annual earnings, 1960 population): Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).

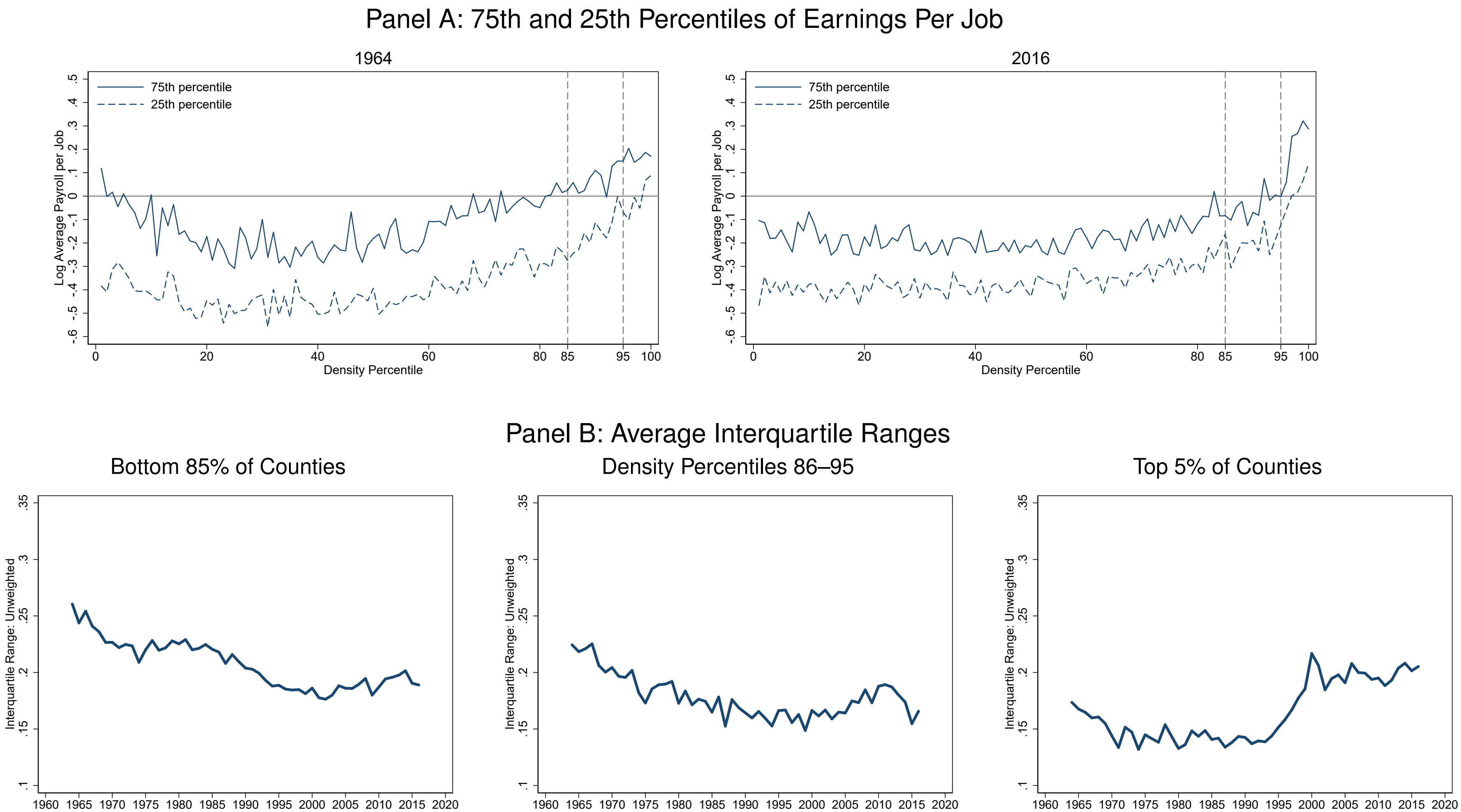


Figure 11. Dispersion in Average Earnings per Job.
Source: US Census Bureau (land area, 1964-2016 annual earnings, 1960 population): Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).

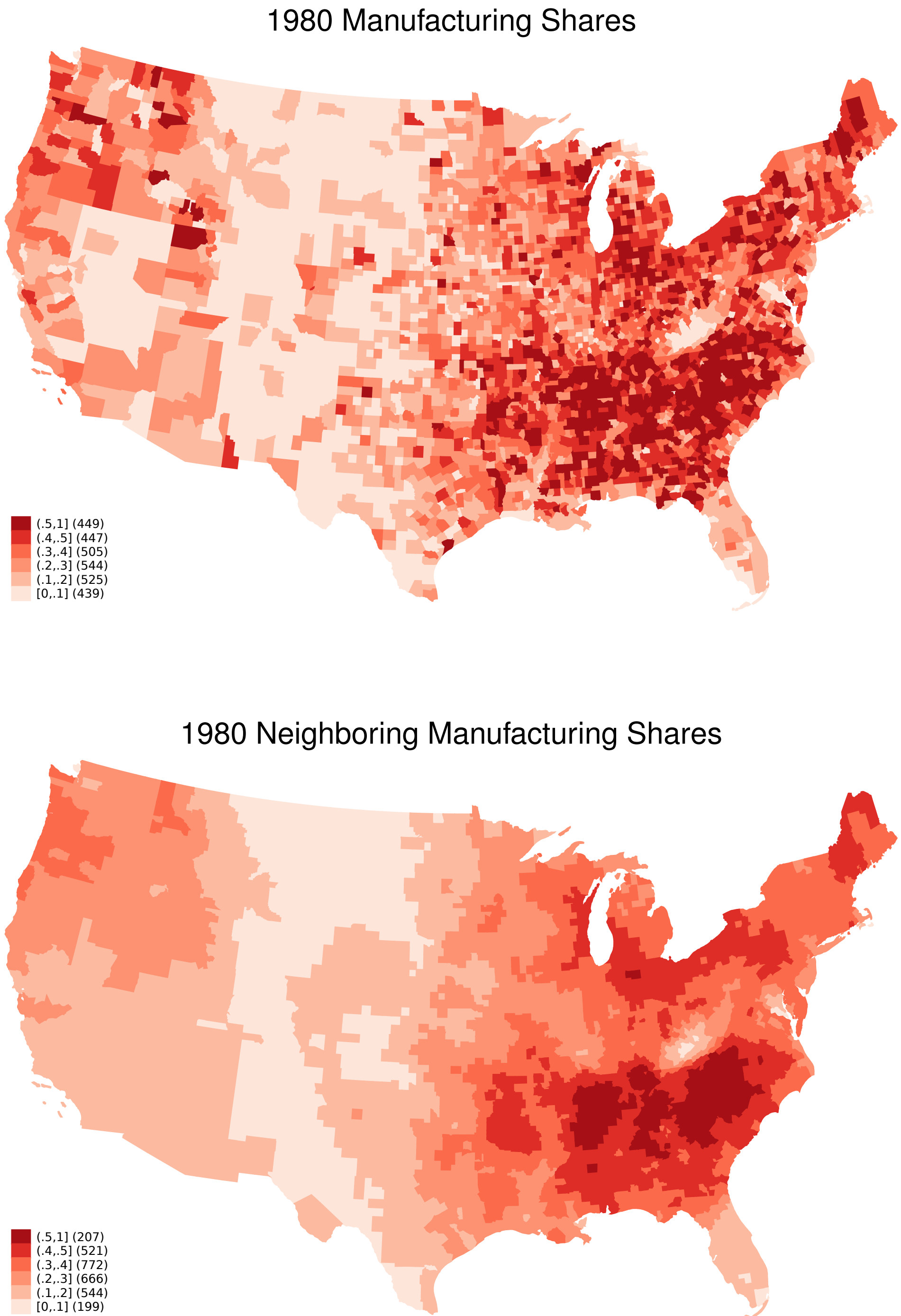


Figure 12. Manufacturing Shares and Neighboring Manufacturing Shares in 1980. *Note:* Neighboring shares are calculated using a second-order contiguity matrix. *Source:* US Census Bureau (shapefiles, 1964-2016 total and manufacturing employment).

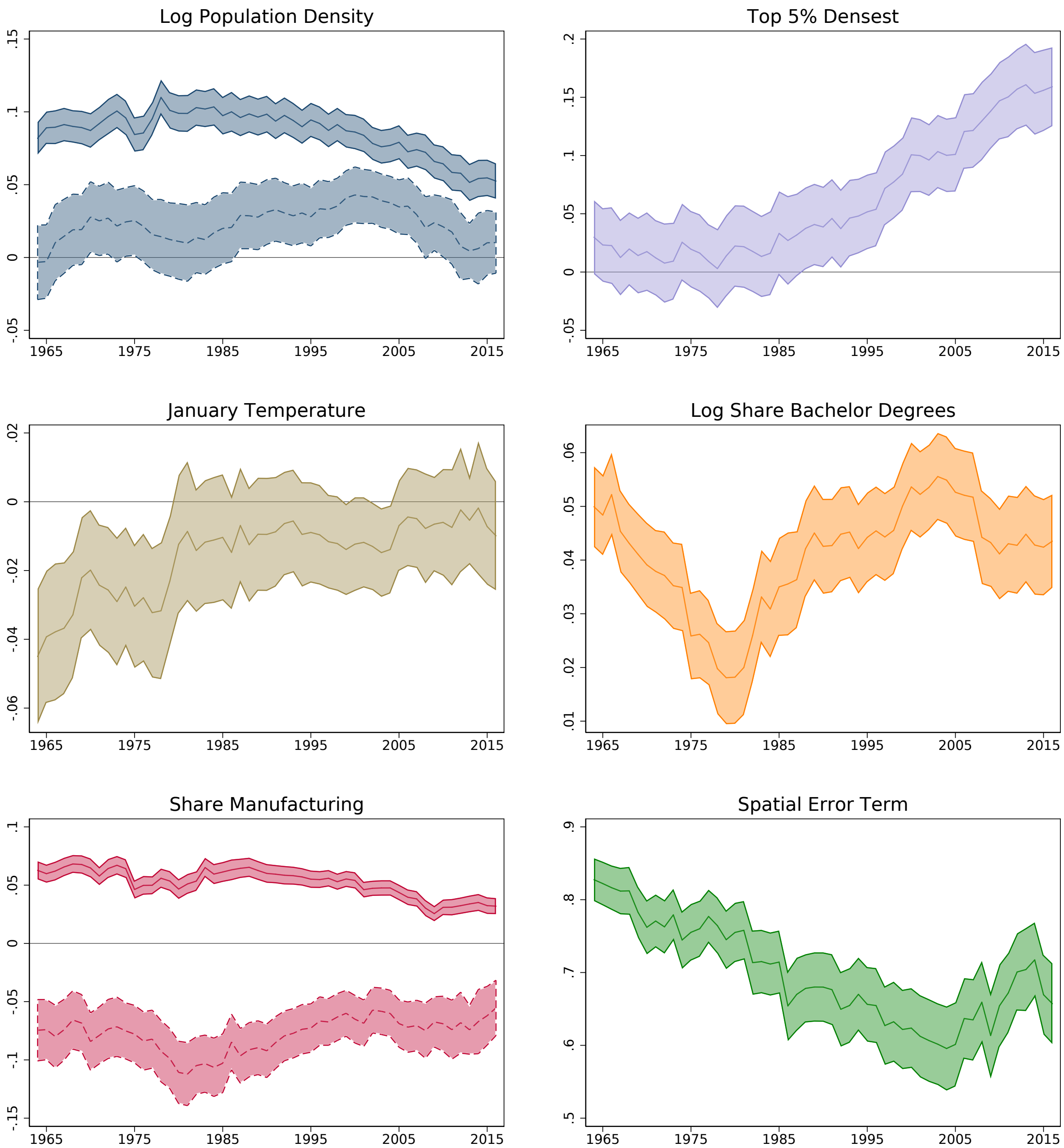


Figure 13. Results of Log Average Earnings Per Job Regressions.
Source: US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), total and manufacturing employment 1964-2016,annual payroll 1964-2016); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).

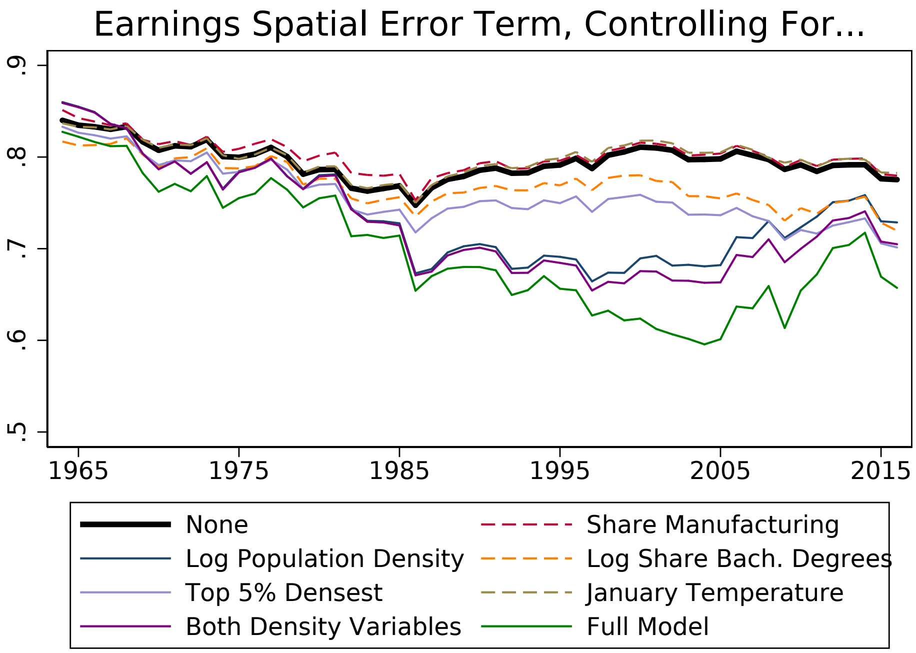


Figure 14. Exploring the Spatial Correlation in County-Level Average Earnings.
Note: Each of the lines above display the importance of the spatial error term (that is, the estimate of λ_ts) in county-level earnings regressions. The green line comes from the baseline model and is thus identical to the spatial error term reported in bottom right panel of Figure 13. The other lines come from models in which only a single regressor, or set of regressors, is included in the model. The dark blue and purple lines come from models that include the log population density term. The figure therefore shows the that decline in the spatial error term in the baseline model indicates a compression of errors *conditional* on county-level density.
Source: US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), total and manufacturing employment 1964-2016, annual payroll 1964-2016); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).

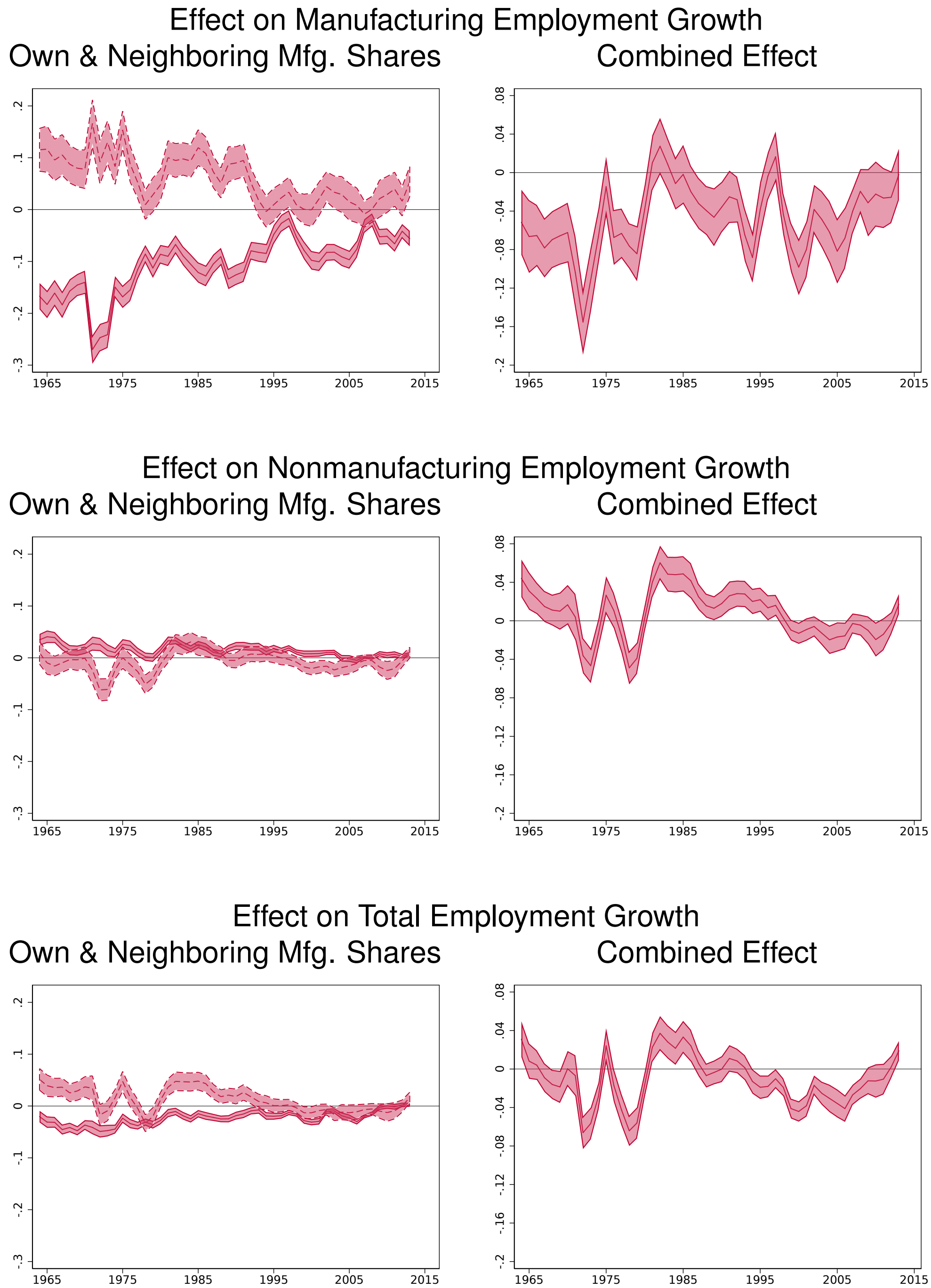
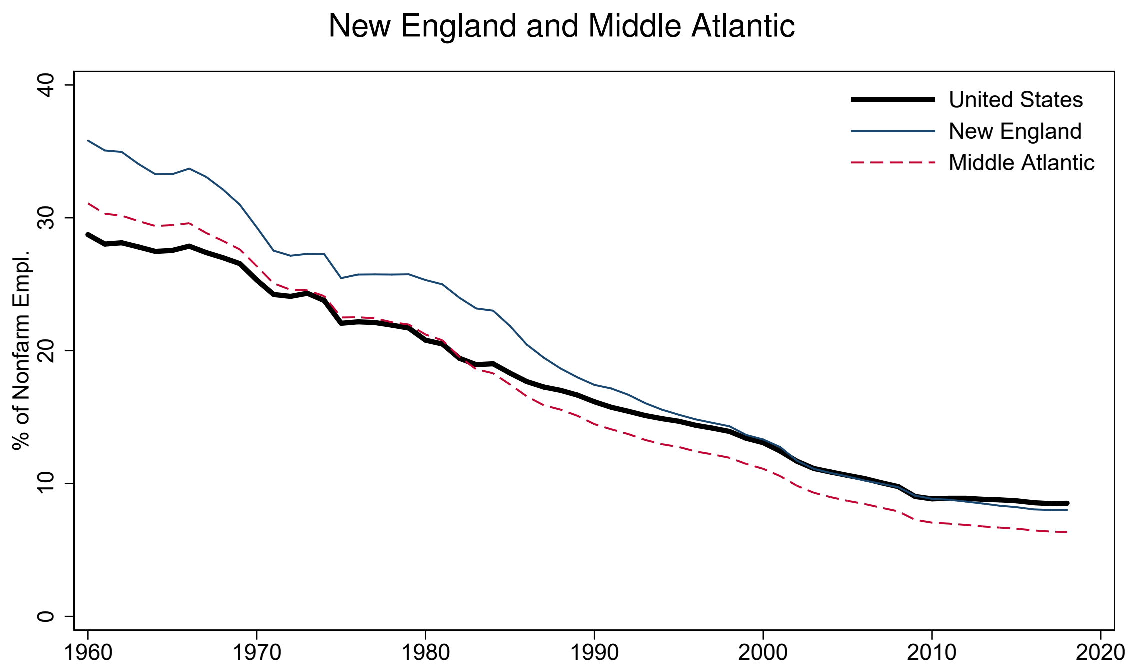


Figure 15. Effect of Manufacturing Shares on Growth of Manufacturing Employment, Nonmanufacturing Employment and Total Employment.

Note: Own-share effects depicted by solid lines; neighboring-share effects depicted by dotted lines. *Source:* US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), total and manufacturing employment 1964-2016); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).



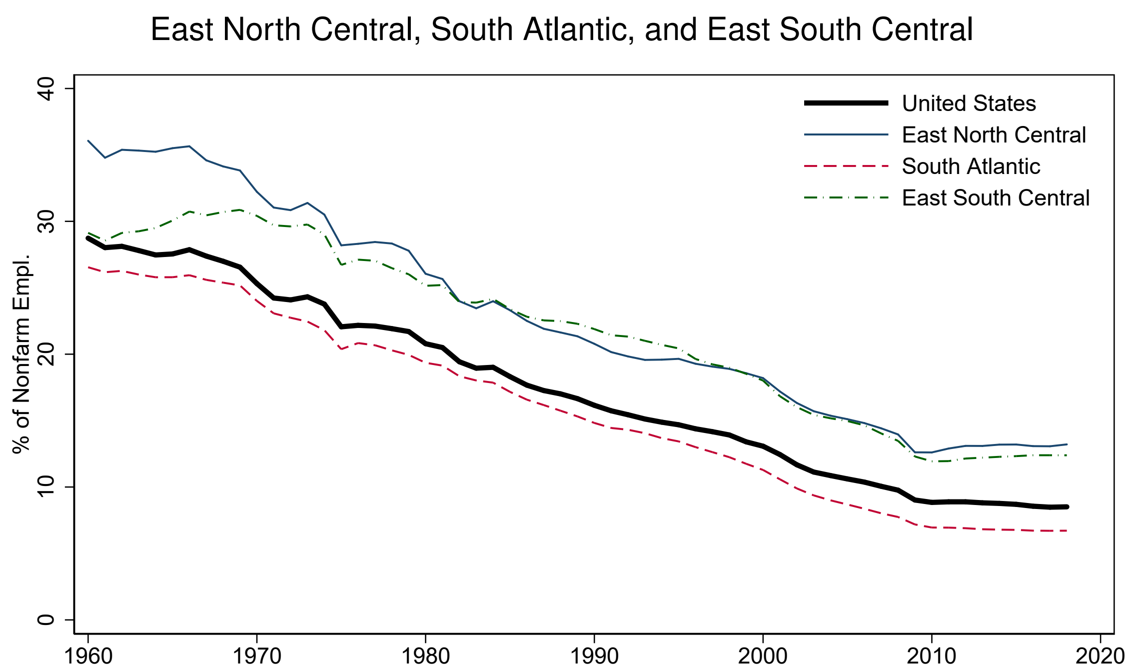


Figure 16. Manufacturing-Employment Shares in Selected Census Divisions.
Note: Census Divisions depicted are: New England (CT, ME, MA, NH, RI, VT), Middle Atlantic (NJ, NY, PA); East North Central (IN, IL, MI, OH, WI); South Atlantic (DE, DC, FL, GA, MD, NC, SC, VA, WV); and East South Central (AL, KY, MS, TN). *Source:* Bureau of Labor Statistics.

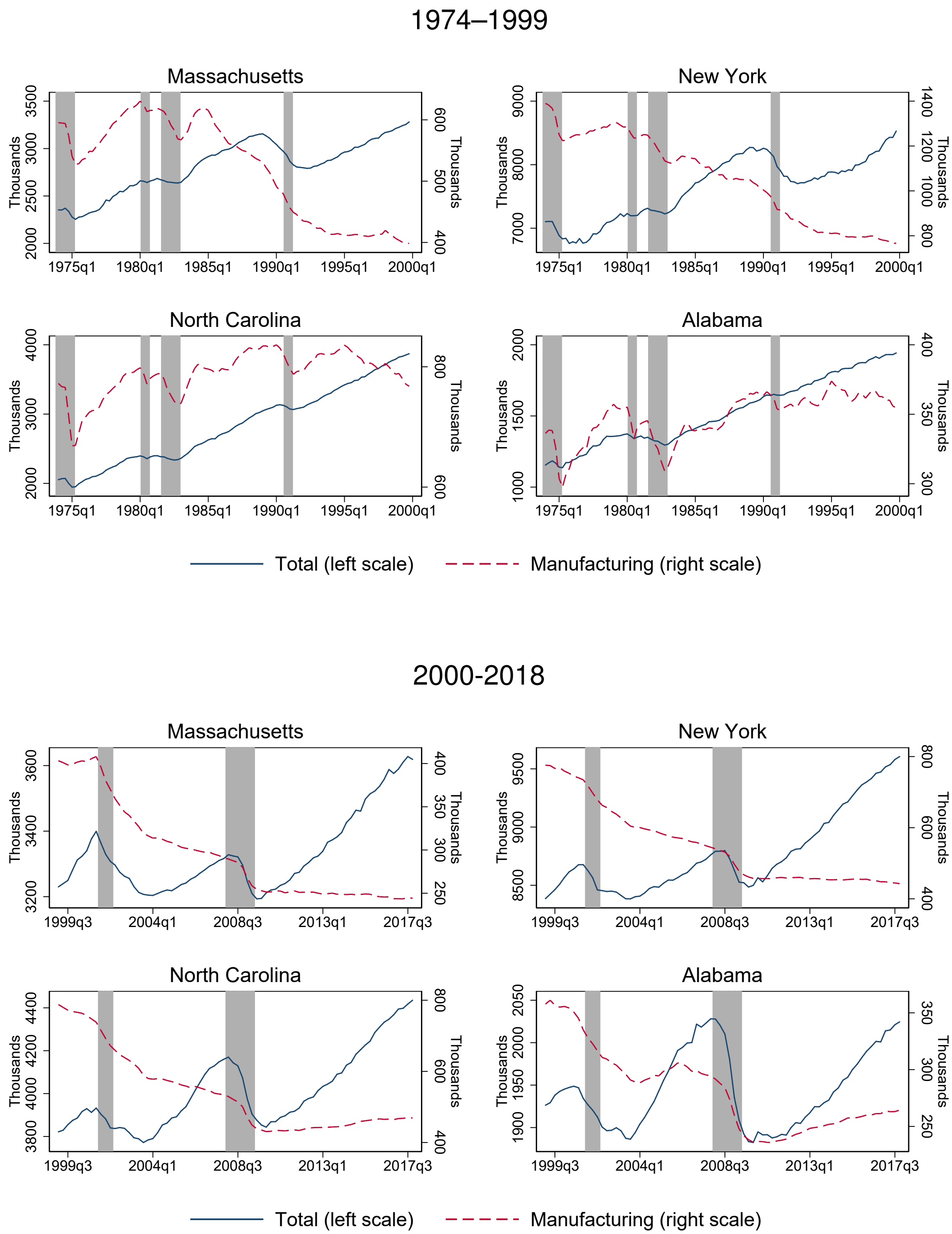
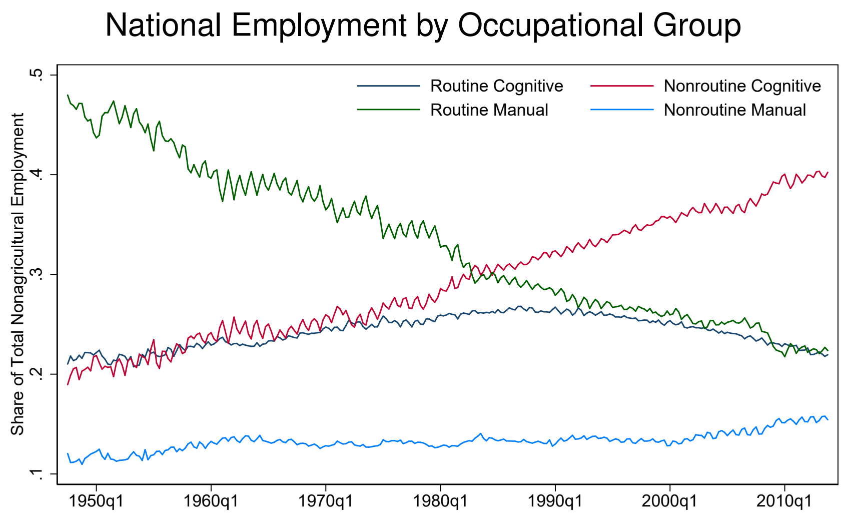


Figure 17. Total and Manufacturing Employment in Four Selected States and Two Time Periods.
Source: Bureau of Labor Statistics.



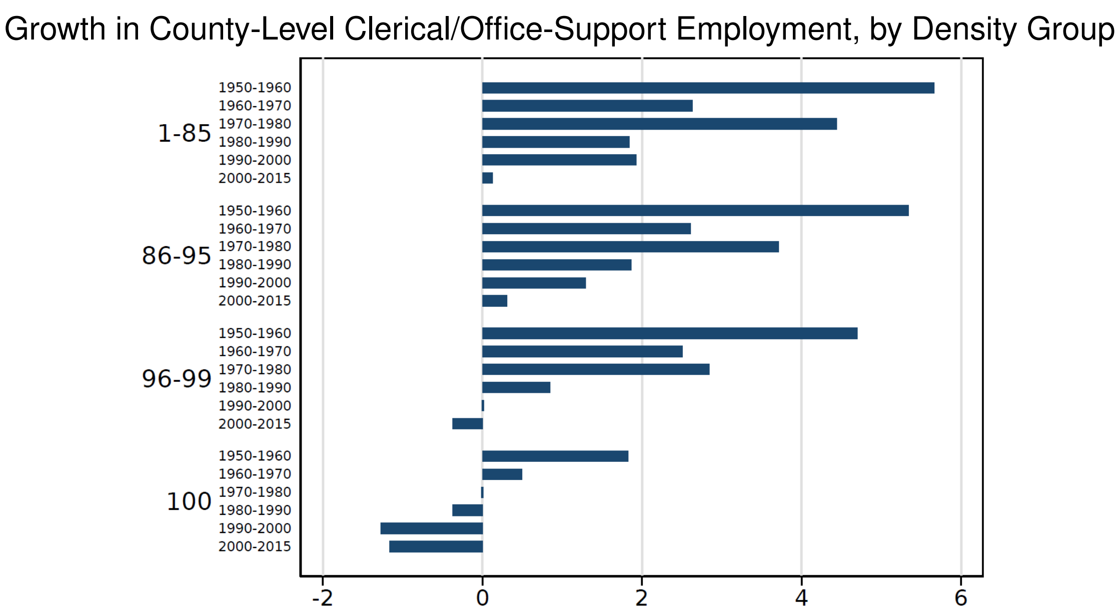


Figure 18. The Evolution of Occupational Employment in the Post-War US, and the Change in Office/Clerical Employment by County Density.
Source: Bureau of Labor Statistics data as analyzed by Foote and Ryan (2014) (top panel) and Bureau of the Census (bottom panel).

A Appendix

A.1 Calculating Manufacturing Employment in the CBP

In the 1964-69 CBP, data are available down to the major group ("one-digit") level, but no finer. In 1970-97, data are available for finer-grained industrial codes, but do not necessarily appear if the count is small; in other words, not every county-industry is reported and establishment counts at a given industrial level do not necessarily add up to the aggregate. Starting in 1998, when CBP switched from the Standard Industrial Classification System (SIC) to the North American Industrial Classification System (NAICS), industrial reporting became exhaustive—that is, 6-digit establishment counts equal aggregate counts and every level in between. As discussed below, we can impute employment for suppressed cells, but we cannot impute cells that do not appear in the data at all. For this reason, we consider only total manufacturing (a major group) in this paper.

One potential problem with using total manufacturing employment in CBP is the change in the definition of manufacturing in the switch from SIC to NAICS. See the comparisons of total employment, manufacturing employment, and the share of manufacturing in employment with statistics from the Bureau of Labor Statistics in Figures A.8, A.9, and A.10 respectively. When we use the share of manufacturing in a county's total employment as a left- or right-hand-side variable in a regression, the resulting measurement error should not matter much since the number of employees shifted in or out of manufacturing would be small relative to a county's total employment. However, when considering manufacturing employment growth, the measurement error may be large relative to the actual manufacturing presence in a county. This means that when we use manufacturing employment growth as a left-hand-side variable, three-year periods that include 1997-98 could have coefficients that are biased towards zero, and the coefficients for these years (1995-1998, 1996-1999, 1997-2000) are indeed much closer to zero than the surrounding years.

Another potential problem with using employment counts in CBP in general is a change in the assignment to counties of employees who were employed by a given establishment but did not physically work there. In 1964-73, the Census Bureau attempted to place employees in the county in which they physically worked. In 1974 onward, these employees were simply assigned to the county where that establishment is located. The impact of this change in methodology can be seen in Figure 5, where the lines for the densest 1% of counties jump by several percentage points in 1973-74. However, this change is small relative to the total change of about 20-25 percentage points in these lines over the entire sample period. Our levels regressions do not feature significant changes in coefficients after this year. The jumps in coefficients around this year in our growth regressions that we attribute to the early 1970s recession are not central to our argument, and the change in methodology cannot explain the similar jump concurrent with the early 1980s recession.

Aggregate county employment counts and one-digit employment counts are suppressed to prevent disclosure in 0.5% and 25% of cases respectively. However, the counts of firms in a county-industry-year by employment range[1] are not suppressed. CBP also provides statistics on a national level,[2] which we use to calculate employment per firm conditional on a firm being in a given employment range[3] by pooling data across 1983-2016.[4] We then use firm counts multiplied by employment per firm to impute employment in the suppressed cells. The strategy of combining firm counts and employment per firm was also used in Glaeser and Kahn (2001) and Autor, Dorn, and Hanson (2013), although their methods for obtaining employment per firm differ.[5]

We make two final adjustments. In the 0.5% of cases where *aggregate* employment is suppressed, this is generally done to prevent disclosure of a particular major industry group via subtraction from the aggregate, while other major groups are not suppressed. In these cases we take the sum of the (imputed and non-imputed) industrial employment in the county rather than the imputation of aggregate employment since this is more accurate. In cases were aggregate county employment is not suppressed but at least two industries are suppressed, we multiply imputed industry employment by a scalar to force the sum of industry employment to equal aggregate employment.[6]

To demonstrate the quality of fit for this imputation strategy, we present two figures. The first shows log employment vs. log of what imputed employment would be for cells that are not imputed (Figure A.12), which demonstrates that we can impute the level of employment with a high degree of accuracy. The second shows the time series of employment for a randomly chosen subset of county-industries that are imputed for some but not all years

[1]In 1974-2016 these ranges are 1-4, 5-9. 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1,000-1,499, 1,500-2,499, 2,500-4,999, and 5,000+.

[2]Specifically, total number of firms and employment in a size class.

[3]CBP pools the 1,000-1,499, 1,500-2,499, 2,500-4,999, and 5,000+ for national statistics. Rather than lose accuracy by using pooled values, we observe that employment per firm closely matches a heuristic: if b_1 and b_2 are the lower and upper bounds of a size class, then employment per firm in the national statistics is very similar to $b_1 + (b_2 + 1 - b_1) \times 0.375$. We also use this heuristic when size classes of 1-3, 4-7, and 8-19 replace 1-4, 5-9. 10-19 in 1964-73. For the largest size class of 5,000+, we obtain employment per firm as the coefficient of a regression of employment on numbers of firms in size classes, where the other coefficients are constrained to employment per firm values that we have already calculated and the constant is constrained to zero. In 1964-73, the largest class size in the county data is 500+, so we assume that firms in this size class have a probabilities of falling into these larger bins that are proportional to the distribution of firms among these bins in that particular county-industry in 1974.

[4]While national data is also available for 1975-1982, employment per firm declines substantially for all size classes in 1982-83, suggesting a change in methodology.

[5]Glaeser and Kahn (2001) use the midpoint of employment ranges, while Autor, Dorn, and Hanson (2013) use a procedure where employment counts are regressed on firm counts, the coefficients are used as estimates of employment per firm, predicted values are used as imputed employment counts, and the procedure is repeated until convergence.

[6]This proportional adjustment is also used in Autor, Dorn, and Hanson (2013).

(Figure A.13), which demonstrates that the resulting changes in the levels are also accurate.

Feasible implementation of repeated cross-section spatial regression with the same weighting matrix requires that there be no missing values in any of the variables.[7] This means that we need to have values for manufacturing employment growth even when manufacturing employment is growing from zero.[8] In these cases, we calculate the growth rate as starting from one employee rather than zero. To mitigate the well-known issues of this approximation, we winsorize this and other growth rates to have a maximum value of 200%. This winsorization affects less than 1.5% of observations.

A.2 Calculating Total County Payrolls

CBP includes annual and first quarter payroll from 1975 on. In 1964-69, there is only annual payroll, and in 1970-74, there is only first quarter payroll. To obtain a series of earnings covering the entire sample period, we use first quarter payroll times four in the place of annual payroll in 1970-74. In 1975-2000, annual payroll is consistently higher than the first quarter times 4, and so there is a jump when switching from the first quarter to annual in 1974-75. However, since we are concerned with earning levels rather than changes, this is unlikely to affect our analysis.

Total payrolls are suppressed in the same 0.5% of county-years that total employment is suppressed. We interpolate the log of per-worker pay and combine this with imputed employment to recover total payrolls. In cases requiring extrapolation (i.e., the first or last n years of data in a county are suppressed), we use a linear time trend in log per-worker pay.[9]

A.3 Other Data

We take annual county-level population and demographic data from the Surveillance, Epidemiology, and End Results (SEER) program run by the National Institutes of Health. Since this data extends back only to 1969, we interpolate 1964-68 using the 1960 Census (Manson et al. 2019). We take data on the number of bachelor degree holders by county from from the 1950, 1970, 1980, 1990, and 2000 Censuses, and 2008-12 and 2013-17 ACS (Manson

[7]We implement these regressions in Stata 15. If there is a missing value for any variable in a spatial regression, then the estimation sample does not match the weighting matrix which must then be re-normalized, which takes a relatively large amount of time.

[8]Because small counties are grouped with larger ones, we never have this issue with aggregate or non-manufacturing employment.

[9]One sparsely-populated county in Texas mainly has zero employment after 2000. For this county, rather than using a time trend for an extended period of time, we use national earnings growth rates and the last non-suppressed value of log per-worker earnings in 1997. Other counties requiring extrapolation have at most 4 missing values.

et al. 2019). In our yearly regressions, we interpolate between these years and set 2016 values equal to 2015 values. We construct seven consistent occupational groups[10] from the same Census an ACS years as we used for bachelor degree holders (Manson et al. 2019). For mapping, geographic contiguity, and land area data, we use the Census Bureau's 2010 county-level Cartographic Boundary File. We take 1941-1970 January temperature from the US Department of Agriculture's natural amenities dataset.

A.4 Geographic Imputation and Consistent County Groupings

All of our county-level data are in terms of consistent county groupings in order to efficiently run repeated cross-section spatial regressions. We use the county groupings supplied by the University of Wisconsin-Madison's Applied Population Lab as part of their net migration data (Winkler et al. 2013). Since our period of study extends before theirs, we supplement their county groupings with retired Virginia FIPS (Federal Information Processing Standard) codes to ensure that dissolved Virginia counties and independent cities are grouped.

In 1950-2015, four counties were created out of pieces of other counties in the lower 48 states.[11] To convert historical count data to modern geographies, we multiply the pre-partition count by the share of the new county's count in the sum of the new county and old county's count in the year immediately after partition and use this as the new county's pre-partition value. We then subtract the new county's value from the old county in the years before partition. In the case of Broomfield County (08-014), which was created out of four counties, it is more accurate to construct values by multiplying the old county's value by the share of transferred population[12] in the old county's population. In the case of advanced degree holders, for which we only have one year of data, we use the share of transferred population in old county population for all four new counties.

In cases where we impute counts that in the original data add up to an overall count (for example, employment by industry adds up to total employment), the sum of the imputations is not necessarily equal to the imputation of the sum. In these cases, we take the imputation of the sum to be the more accurate value and multiply all categories in the county-year by a scalar to satisfy the adding-up constraint.

Lastly, there are some small counties in the data that experience very large growth

[10] See Table A.1 for details on the composition of these groups and Figure A.14 to see their aggregate levels over time.

[11] Menominee, WI (55-078) from Shawano, WI (55-115) in 1959; Cibola, NM (35-006) from Valencia, NM (35-061) in 1981; La Paz, AZ (04-012) from Yuma, AZ (04-027) in 1983; and Broomfield, CO (08-014) from Adams (08-001), Boulder (08-013), Jefferson (08-059), and Weld, CO (08-123) in 2001. While Menominee was created before the beginning of our analysis, it does not appear in the 1960 Census data used to interpolate population in 1964-68.

[12] This information is available on the Census Bureau's county changes webpage: https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.2000.html

rates in small periods of time. To prevent these small counties from biasing our results, we combine counties whose minimum employment over the sample period fall below 200 with neighboring counties, taking care to respect the boundaries of commuting zones and Core-Based Statistical Areas (CBSAs) where possible. When not possible, we ensure that the receiving commuting zone is as small as possible[13] and that the receiving CBSA is a micropolitan (i.e., small metro) area rather than a metropolitan (i.e., large metro) area. After this procedure is completed, we are left with 2,909 consistent county groupings covering the entire contiguous United States.

[13] Many of these small counties are in a single-county commuting zone, so it is better to combine them with another small commuting zone rather than a large one, whose boundaries we most wish to respect.

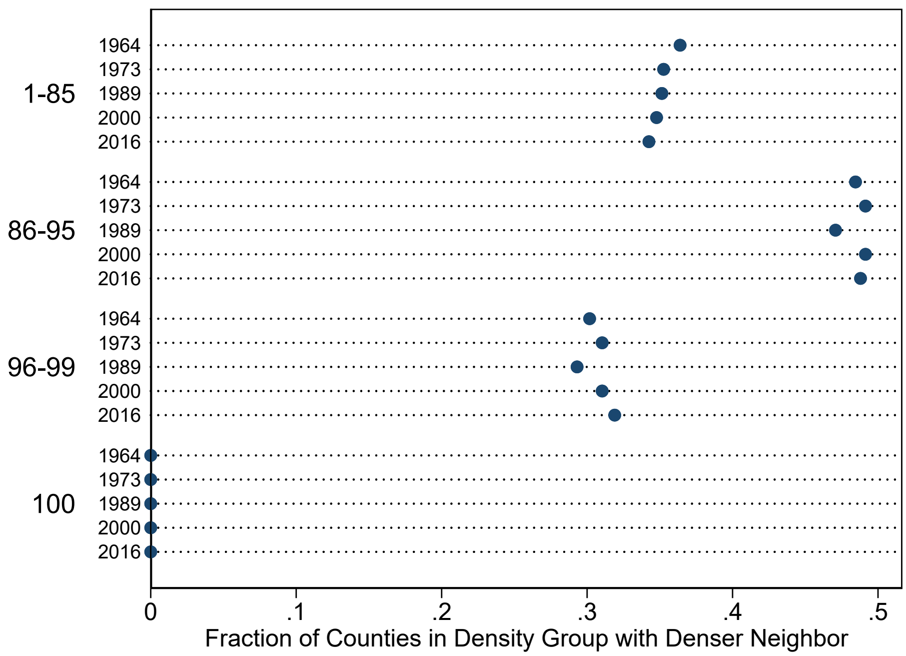
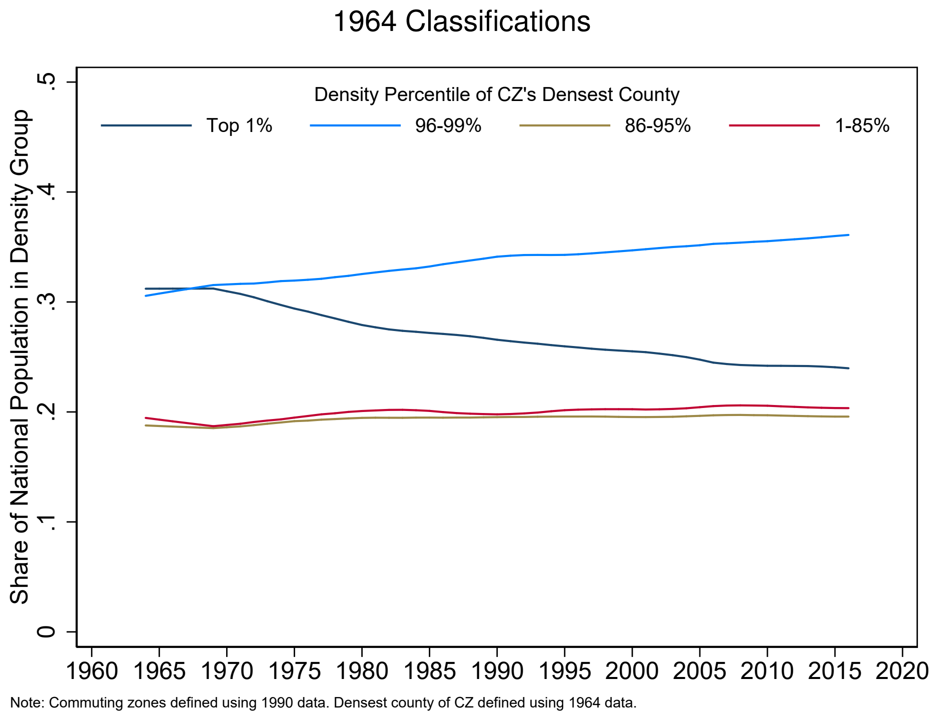


Figure A.1. Counties in the Four Density Groups with Neighbors in a Denser Group. *Source:* US Census Bureau (land area, 1960 population, shapefiles); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population).



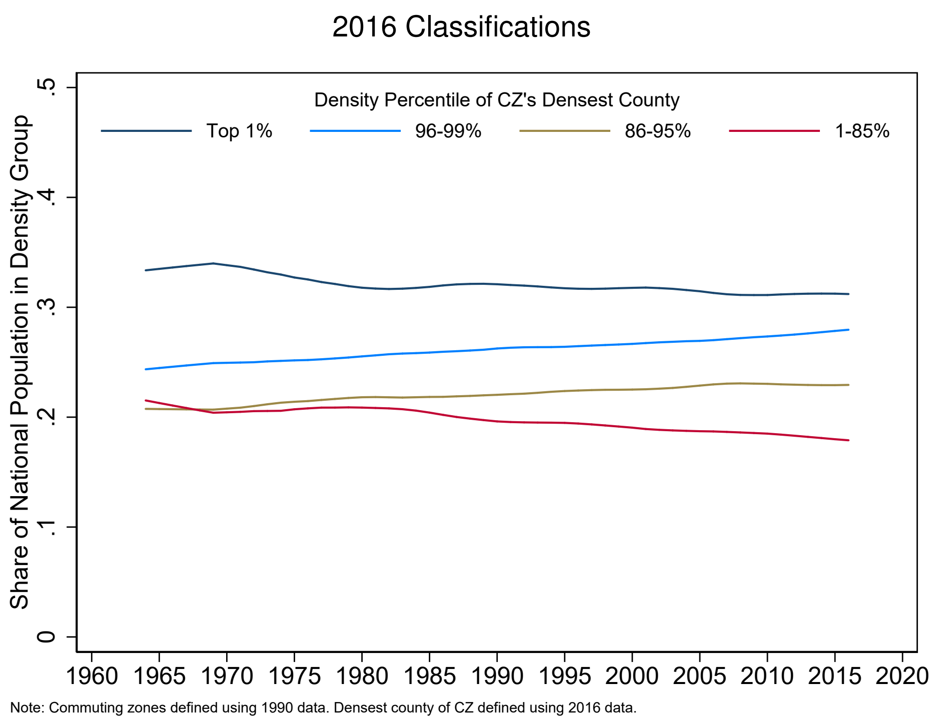


Figure A.2. Population Shares within Commuting-Zone Groups: 1964–2016.
Source: US Census Bureau (land area, 1960 population); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (1990 commuting zone definitions).

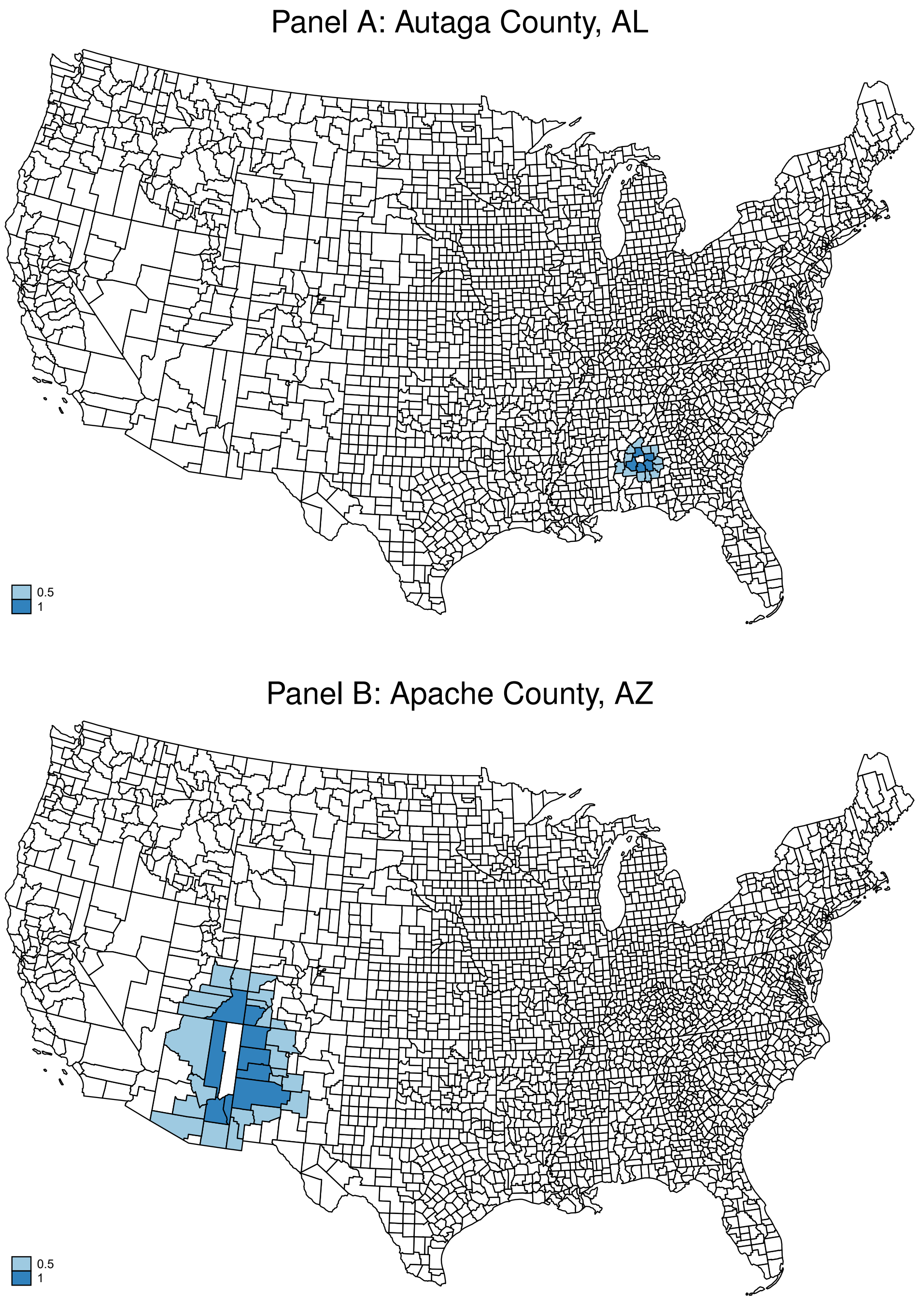


Figure A.3. Illustrating the Second-Order Contiguity Matrix with Two Counties. *Source:* US Census Bureau (shapefiles).

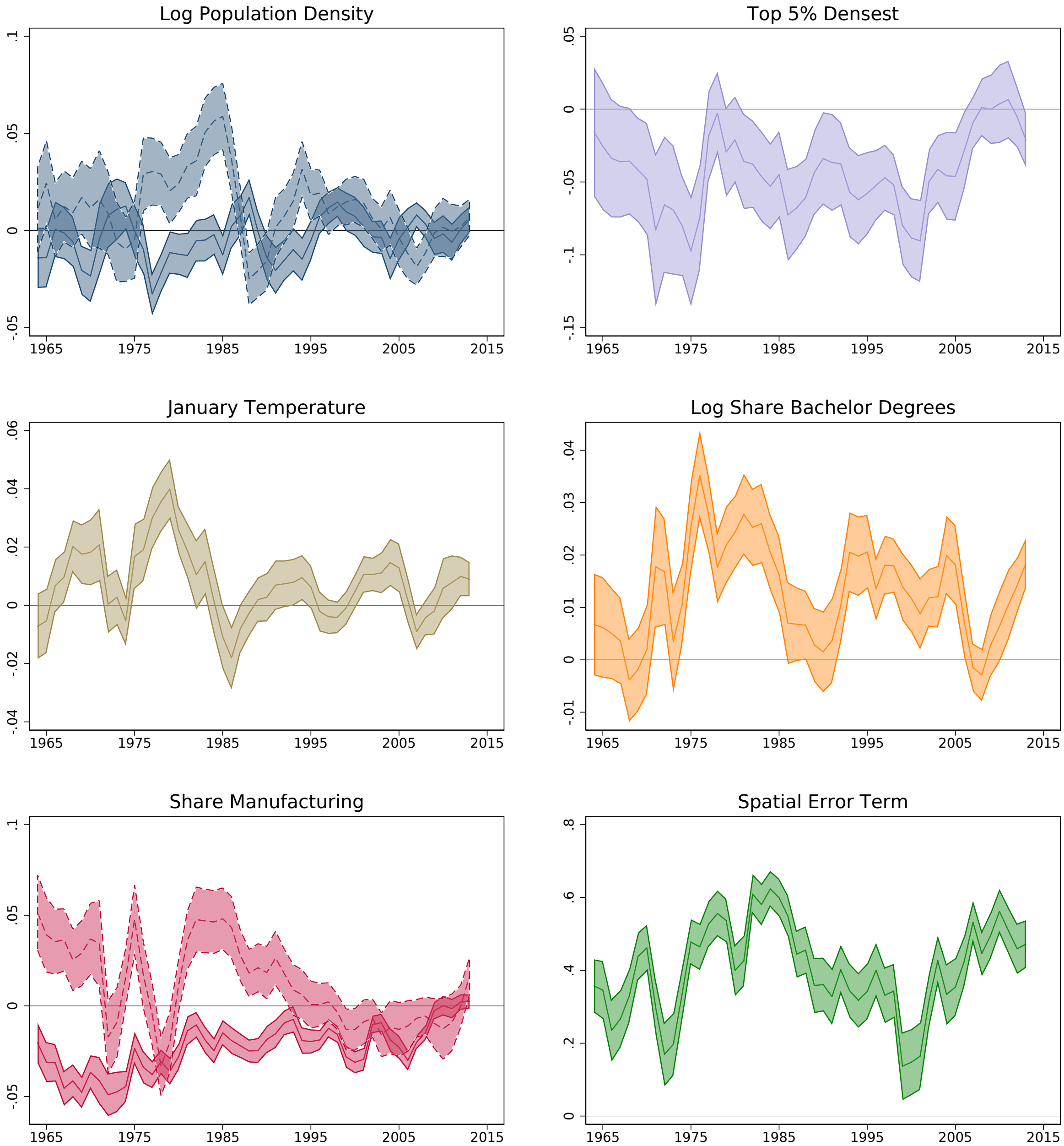


Figure A.4. Results of 3-Year Total Employment Growth Regressions.
Source: US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), total and manufacturing employment); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).

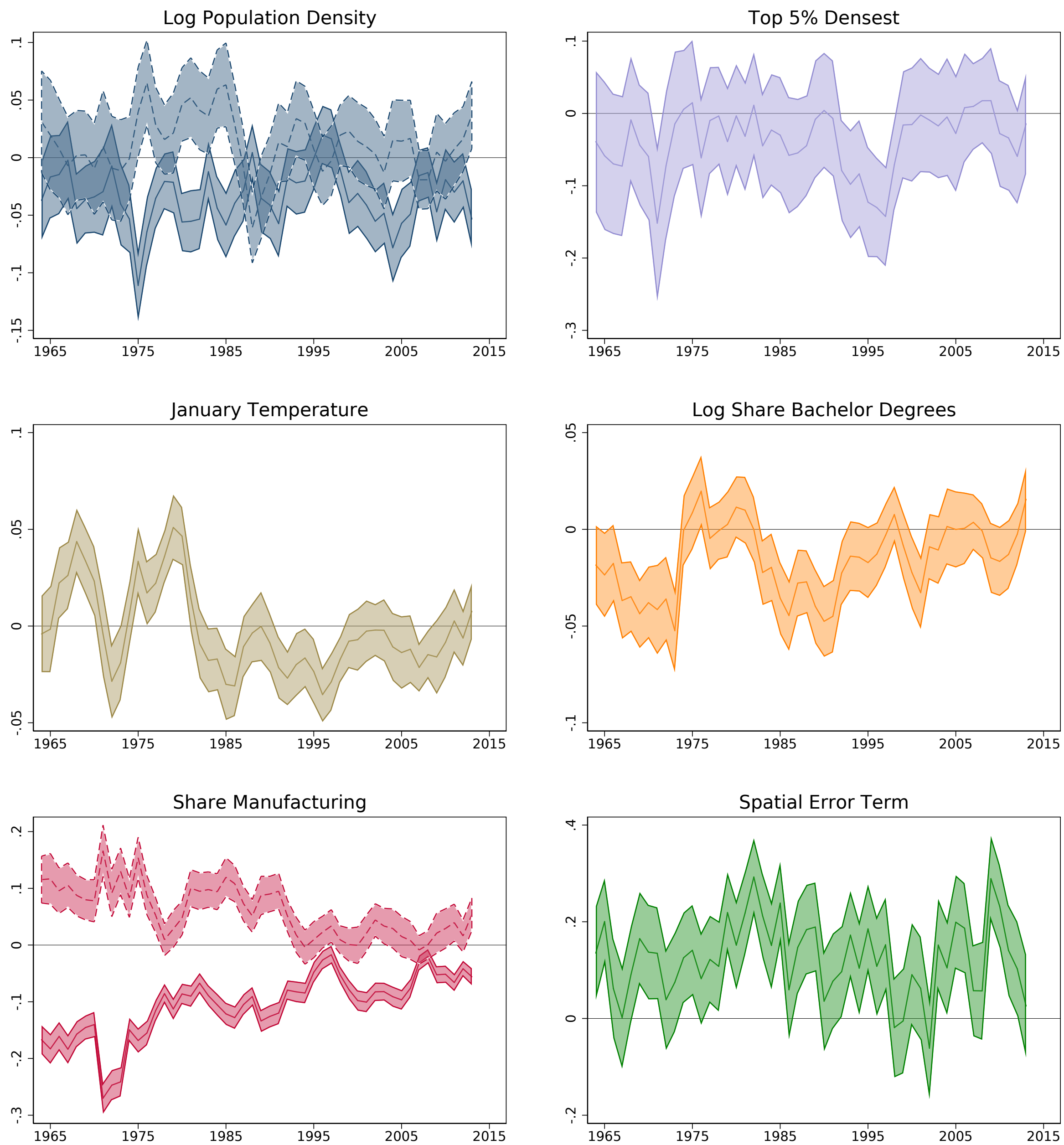


Figure A.5. Results of 3-Year Manufacturing Employment Growth Regressions.
Source: US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), total and manufacturing employment); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).

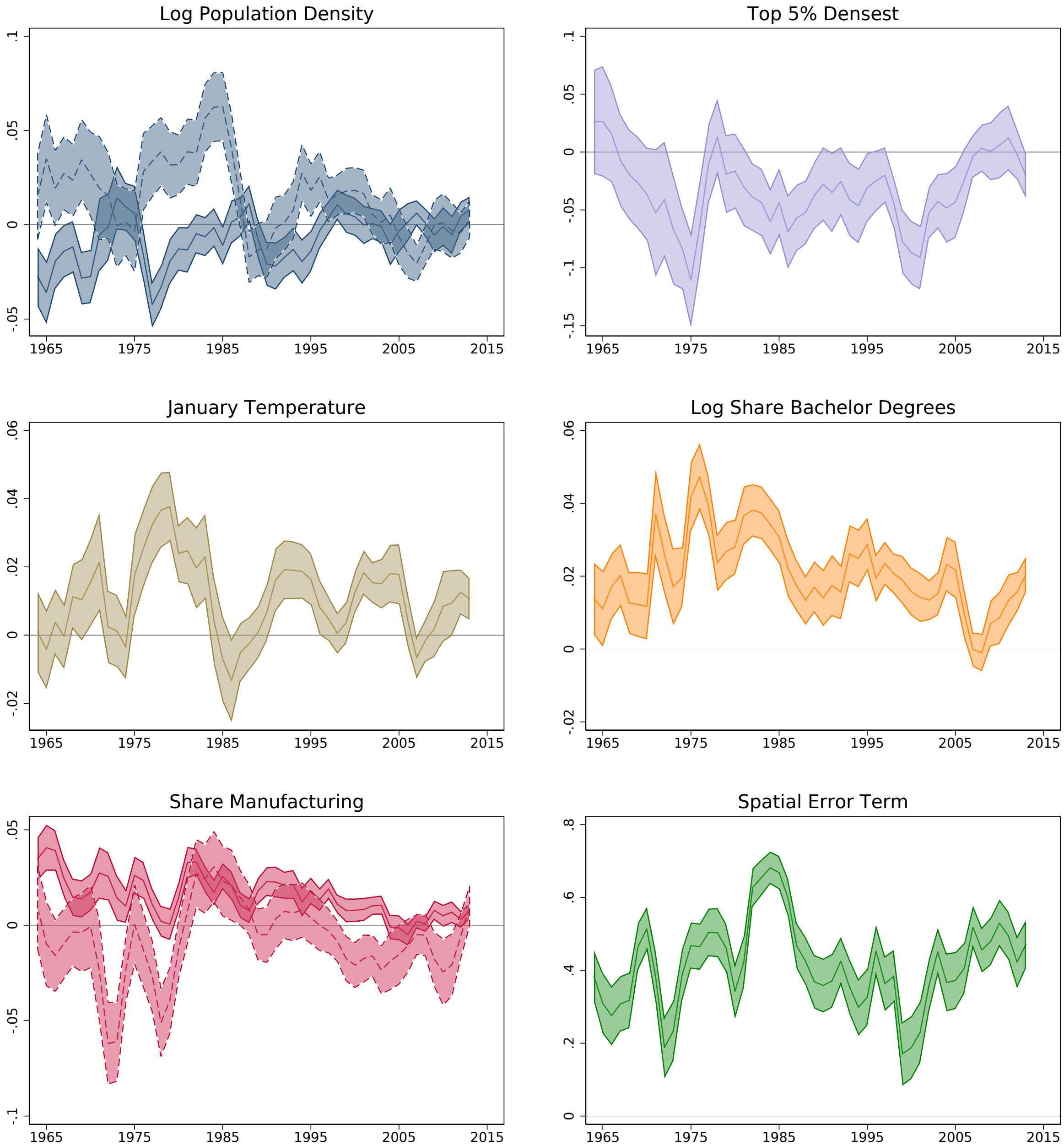


Figure A.6. Results of 3-Year Nonmanufacturing Employment Growth Regressions.
Source: US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), total and manufacturing employment); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature).

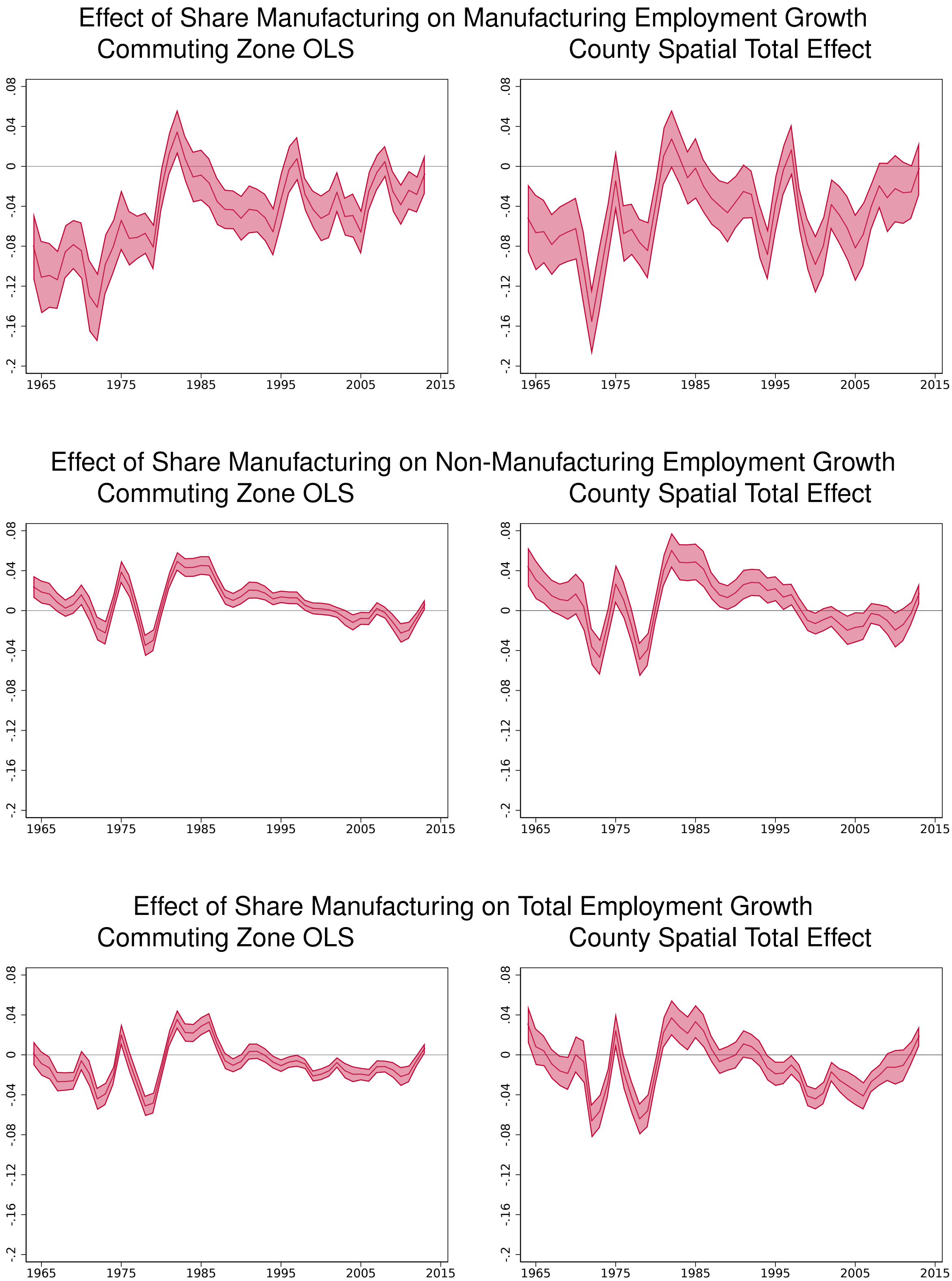


Figure A.7. Comparing Job Growth Results: County-Level Spatial Regression Total Effects Vs. Commuting Zone-Level OLS Coefficients .
Source: US Census Bureau (land area, 1960 population, shapefiles, share bachelor's degrees (1950, 1970, 1980, 1990, 2000, 2010, 2015), 1964-2016 total and manufacturing employment); Surveillance, Epidemiology, and End Results (SEER) program (1969-2016 population); US Department of Agriculture (January temperature, 1990 commuting zone definitions).

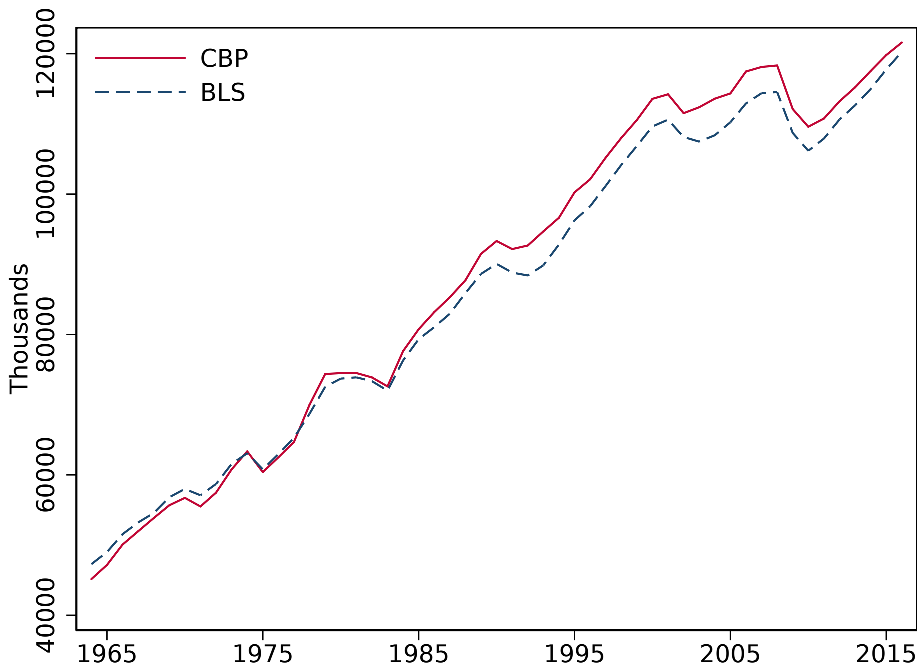


Figure A.8. Total Employment in CBP Vs. Bureau of Labor Statistics
Source: US Census Bureau, Bureau of Labor Statistics.

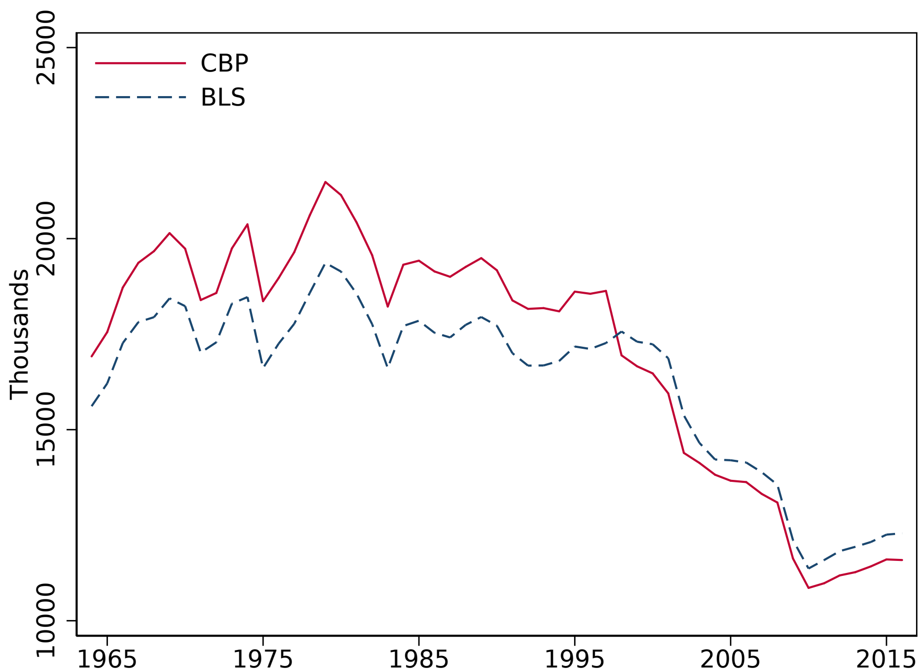


Figure A.9. Manufacturing Employment in CBP Vs. Bureau of Labor Statistics
Source: US Census Bureau, Bureau of Labor Statistics.

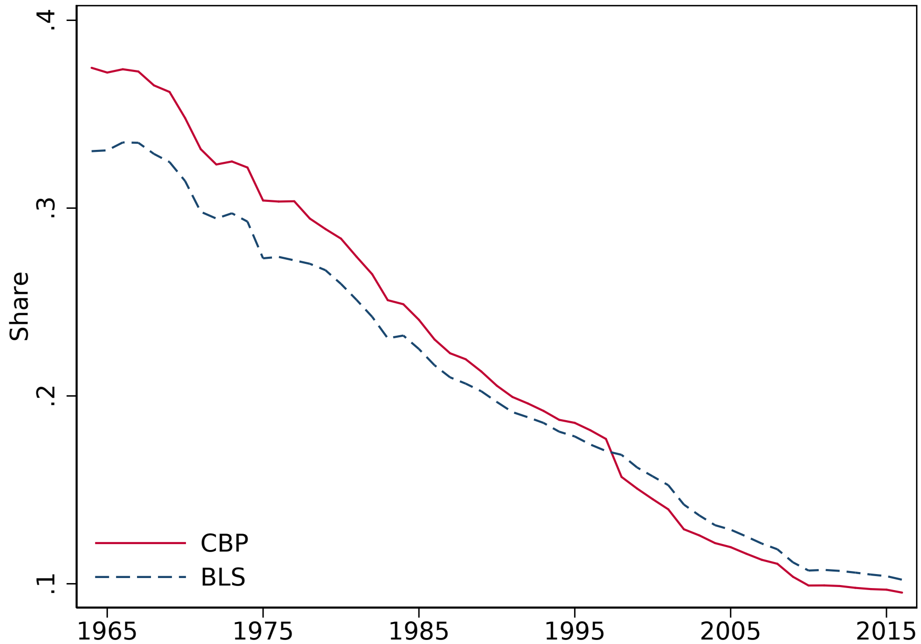


Figure A.10. Share Manufacturing Employment in CBP Vs. Bureau of Labor Statistics
Source: US Census Bureau, Bureau of Labor Statistics.

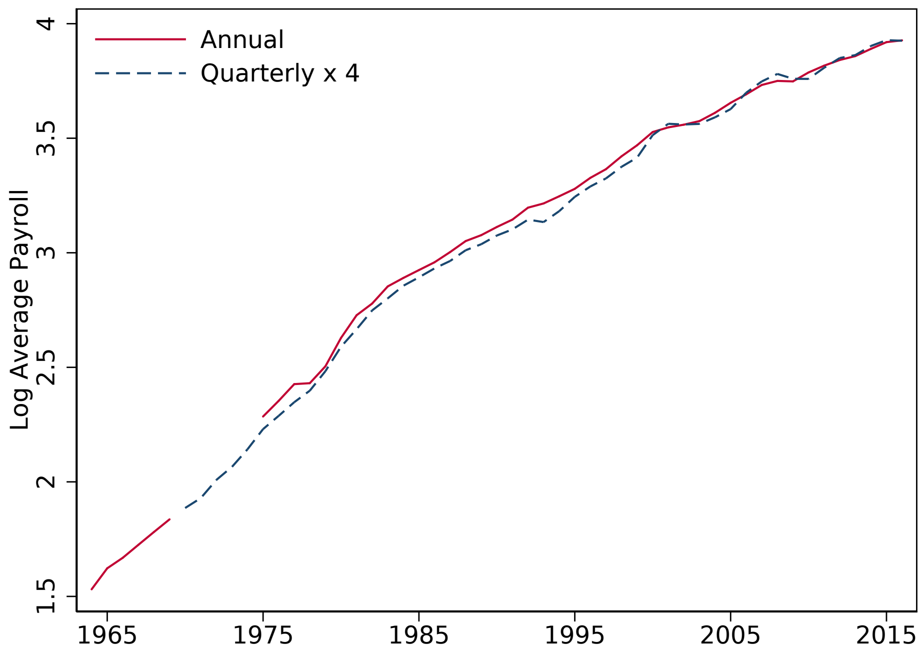


Figure A.11. CBP Payroll Variables
Source: US Census Bureau.

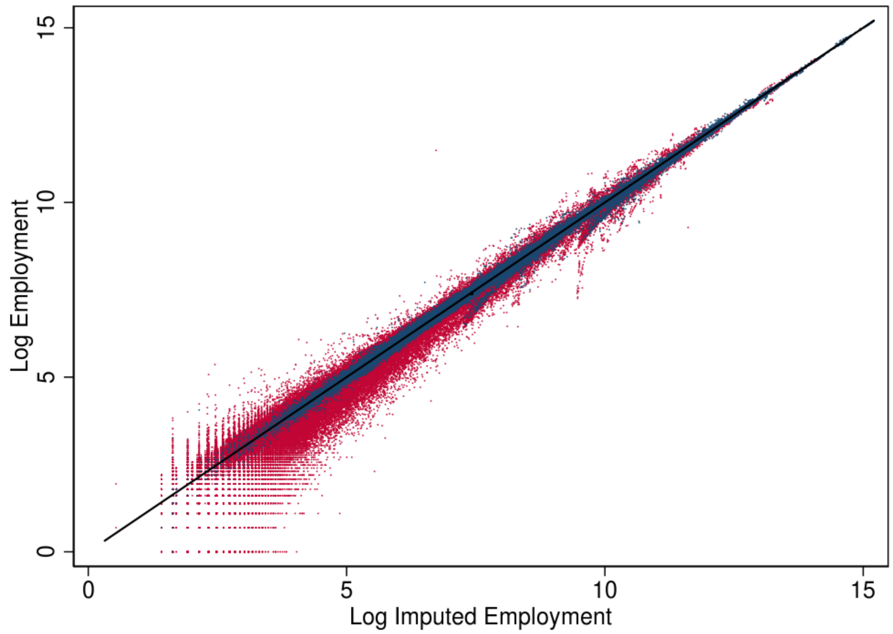


Figure A.12. Actual Employment Vs. What Imputation Would Be, Non-Imputed Aggregates (Blue, $R^2 = 0.9986$) And Major Groups (Red, $R^2 = 0.9891$)

Source: US Census Bureau

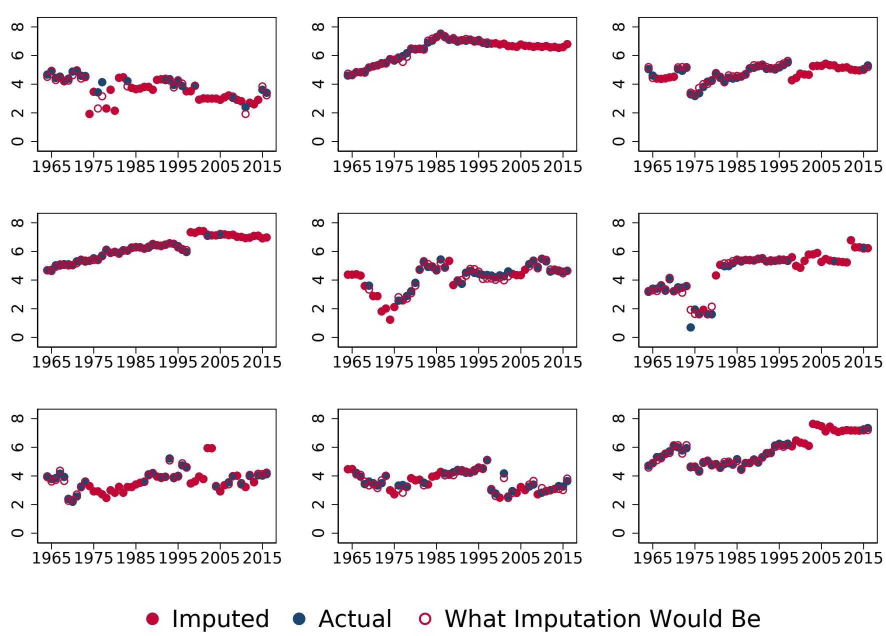


Figure A.13. Time Series View of Log Imputation Fit: Randomly Chosen, Partially Imputed County-Industries

Source: US Census Bureau.

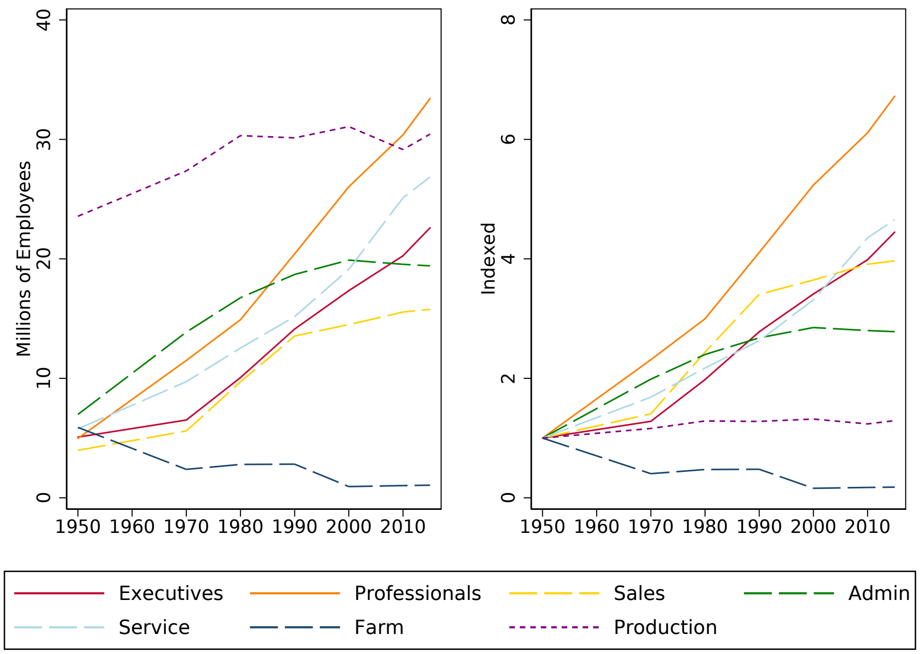


Figure A.14. Consistent Occupation Groups Over Time
Source: US Census Bureau.

	1950	1970	1980	1990	2000	2010	2015
Executives	Managers, officials and proprietors (except farm)	Managers and administrators, except farm	Executive, administrative, and managerial occupations (codes 3-37);	Executive, administrative, and managerial occupations (000-042)	Management, business, and financial operations occupations	Management, business, and financial occupations	Management, business, and financial occupations
Professionals	Professional, technical and kindred workers	Professional, technical and kindred workers	Professional speciality occupations (codes 43-199); Technicians and related support occupations (codes 203-235)	Professional specialty occupations (043-202); Technicians and related support occupations (203-242)	Professional and related occupations	Computer, engineering, and science occupations; Education, legal, community service, arts, and media occupations; Healthcare practitioners and technical occupations	Computer, engineering, and science occupations; Education, legal, community service, arts, and media occupations; Healthcare practitioners and technical occupations
Sales	Sales workers	Sales workers	Sales occupations (codes 243-285)	Sales occupations (243-302)	Sales and related occupations	Sales and related occupations	Sales and related occupations
Admin	Clerical and kindred workers	Clerical and kindred workers	Administrative support occupations, including clerical (codes 303-389)	Administrative support occupations, including clerical (303-402)	Office and administrative support occupations	Office and administrative support occupations	Office and administrative support occupations
Service	Private household workers; Service workers (except private household)	Service workers, except private household; Private household workers	Private household occupations (codes 403-407); Protective service occupations (codes 413-427); Service occupations, except household and protective (codes 433-469)	Private household occupations (403-412); Protective service occupations (413-432); Service occupations, except protective and household (433-472)	Healthcare support occupations; Protective service occupations; Food preparation and serving related occupations; Building and grounds cleaning and maintenance occupations; Personal care and service occupations	Service occupations	Service occupations
Farm	Farm and farm managers; Laborers (except unpaid) and farm foremen	Farmers and farm managers; Farm laborers or foremen	Farming, forestry and fishing occupations (codes 473-499)	Farming, forestry, and fishing occupations (473-502)	Agricultural workers, including supervisors; Fishing, hunting, and forestry occupations	Farming, fishing, and forestry occupations	Farming, fishing, and forestry occupations
Production	Craftsmen, foremen, and kindred workers; Operatives and kindred workers; Laborers, unpaid family workers	Craftsmen, foremen and kindred workers; Operatives, except transport; Transport, equipment operatives	Precision production, craft and repair occupations (codes 503-699); Machine operators, assemblers and inspectors (codes 703-799); Transportation and material moving occupations (codes 803-859); Handlers, equipment cleaners, helpers and laborers (codes 863-889)	Precision production, craft, and repair occupations (503-702); Machine operators, assemblers, and inspectors (703-802); Transportation and material moving occupations (803-863); Handlers, equipment cleaners, helpers, and laborers (864-902)	Construction and extraction occupations; Installation, maintenance, and repair occupations; Production occupations; Transportation and material moving occupations	Construction and extraction occupations; Installation, maintenance, and repair occupations; Production, transportation, and material moving occupations	Construction and extraction occupations; Installation, maintenance, and repair occupations; Production, transportation, and material moving occupations

Table A.1. Consistent Occupational Definitions
Source: US Census Bureau. *Note:* Descriptions are given by Manson et al. (2019).