Cities, Skills, and Regional Change*

Edward L. Glaeser
Harvard University and NBER

Giacomo A. M. Ponzetto
CREI and Universitat Pompeu Fabra

Kristina Tobio
Harvard University

March 31, 2011

Abstract

One approach to urban areas emphasizes the existence of certain immutable relationships, such as Zipf's or Gibrat's Law. An alternative view is that urban change reflects individual responses to changing tastes or technologies. This paper examines almost 200 years of regional change in the U.S. and finds that few, if any, growth relationships remain constant, including Gibrat's Law. Education does a reasonable job of explaining urban resilience in recent decades, but does not seem to predict county growth a century ago. After reviewing this evidence, we present and estimate a simple model of regional change, where education increases the level of entrepreneurship. Human capital spillovers occur at the city level because skilled workers produce more product varieties and thereby increase labor demand. We find that skills are associated with growth in productivity or entrepreneurship, not with growth in quality of life, at least outside of the West. We also find that skills seem to have depressed housing supply growth in the West, but not in other regions, which supports the view that educated residents in that region have fought for tougher land-use controls. We also present evidence that skills have had a disproportionately large impact on unemployment during the current recession.

*Comments are appreciated and can be sent to eglaeser@harvard.edu, gponzetto@crei.cat, and kristina_tobio@ksg.harvard.edu. Financial support from the Taubman Center for State and Local Government, the Spanish Ministry of Science and Innovation (grants ECO2008-01666, CONSOLIDER-Ingenio 2010 CSD2006-00016, and Juan de la Cierva JCI-2010-08414), the Barcelona GSE Research Network and the Generalitat de Catalunya (grant 2009 SGR 1157) is gratefully acknowledged.

1 Introduction

Are there universal laws of urban and regional population growth that hold over centuries, or do time-specific shifts in tastes and technology drive the shifts of population over space? Is urban change better understood with the tools of physics or a knowledge of history? In this paper, we investigate patterns of population and income change over the long run in the older regions of the U.S. Within this large land mass, there has been remarkable persistence in population levels across time. The logarithm of county population in 2000 rises almost perfectly one-for-one with the logarithm of population in 1860 and the correlation between the two variables is 66 percent.

Formal modeling of city growth has naturally tended to focus on patterns that are presumed to hold universally, such as Gibrat's law, which claims that population growth rates are independent of initial levels. Gibrat's law has received a great deal of recent interest because of its connection with Zipf's law, the claim that the size distribution of cities in most countries is well approximated by a Pareto distribution (Gabaix 1999, Gabaix and Ioannides 2004, Eeckhout, 2004).[1] Our paper is not concerned with static laws of urban size, such as Zipf's law, but rather with the permanence of dynamic relationships.

The long-run persistence of county level populations implies that Gibrat's law has very much held in the long run. But Gibrat's law doesn't hold reliably for county population changes at higher frequencies. Before 1860 and after 1970, less populous counties grew more quickly. During the intervening decades, when America industrialized and sectors concentrated to exploit returns to scale (Kim 2006), population growth was regularly faster in more populated areas. One interpretation is that Gibrat's law is universal, but only over sufficiently long time periods. An equally plausible interpretation is that Gibrat's law holds in the long run because of the accidental balancing of centripetal forces, which dominated during the industrial era, and centrifugal forces, which have become more powerful in the age of the car and the truck; and that—as a result—there is no reason to expect the law to hold in the future.

Geographic variables also wax and wane in importance. During recent decades, January temperature has been a reliable predictor of urban growth, and that was also true in the late 19th century; but it wasn't true either before 1860 or in the early decades of the 20th century. The Great Lakes seem to have attracted population both in the early years of the American Republic, and also during a second wave of growth in the first half of the

[1] Another strand of the literature has expanded the standard theory of endogenous growth to incorporate urban dynamics and reconcile increasing returns at the local level with constant returns and a balanced growth path for the aggregate economy (Eaton and Eckstein 1997; Black and Henderson 1999; Duranton 2006, 2007; Rossi-Hansberg and Wright 2007).

20th century, associated with the expansion of industrial cities that formed around earlier commercial hubs. Population has moved away from these waterways since 1970, even within the eastern areas of the U.S. To us, these patterns seem to suggest waves of broad regional change that are associated with tectonic shifts in the economy, rather than time-invariant laws.

Even schooling has its limits as a predictor of growth. Since 1940, in our sample of counties, the share of a county's population with college degrees at the start of a decade predicts population growth in every subsequent decade except the 1970s. Even in the 1970s, schooling predicts growth among counties with more than 100,000 people. But this fact does not hold in the West even today, and it doesn't seem to hold during much of the 19th century. While Simon and Nardinelli (2002) document a connection between skilled occupations and area growth since 1880, we don't find much of a relationship between the share of the population with college degrees in 1940 and growth before 1900. Perhaps this just reflects the fact that we are forced to use an ex post measure of education that may well be poorly correlated with skills in 1860 or 1880; but it seems as likely that the industrializing forces of the late 19th century just didn't favor better educated areas.

The one persistent truth about population change in this group of counties is that growth strongly persists. With the exception of a single decade (the 1870s), the correlation between population growth in one decade and the lagged value of that variable is never less than .3 and typically closer to .5. Among counties with more than 50,000 people, the correlation between current and lagged population growth is never less than .4 in any decade. Over longer seventy-year time periods, however, faster growth in an early period is associated with lower subsequent growth. These facts are quite compatible with the view that growth is driven by epoch-specific forces, like large-scale industrialization and the move to car-based living, that eventually dissipate.

We only have county income data since 1950, and as a result we have little ability to observe large historic shifts in this variable. In every decade except the 1980s there is strong mean reversion in this variable; Barro and Sala-i-Martin (1991) established mean reversion for state incomes going back to 1840. The connection between income growth and education or manufacturing has, however, varied from decade to decade. In the 1960s and 1970s, income growth was positively correlated with income growth during the previous decade, but that trend reversed after 1980. With the exception of mean reversion, universal laws about income growth seem no more common than universal laws about population growth.

One interpretation of the collection of facts assembled in Section 2 is that the eastern United States has experienced three distinct epochs. In the first 60-odd years of the 19th century, the population spread out, especially towards colder areas with good soil quality and access to waterways. From the late 19th century until the 1950s, America industrialized and the population clustered more closely together, which set off a second growth spurt of the Great Lakes region. Over the past four decades, declining transport costs has led both to the spread of people across space, towards the Sun Belt, and the increasing success of skilled, entrepreneurial areas that thrive by producing new ideas. The early period of spatial concentration of U.S. manufacturing at the beginning of the 20th century and its dispersion in the last few decades are quite compatible with the work of Desmet and Rossi-Hansberg (2009a, 2009b, 2010), who suggest that innovative new industries cluster to benefit from knowledge spillovers while mature sectors spread out following technology diffusion.

After reviewing these stylized facts, in Section 3 we present a model of human capital, entrepreneurship and urban reinvention. The model is meant to help us understand the strong connection between human capital and urban reinvention in the post-war period. The model suggests that the impact of skills on growth will differ depending on local conditions, and skills will be particularly valuable in places that are hit with adverse shocks. The model also suggests a decomposition that enables us to understand the channels through which human capital impacts on growth.

Skilled cities may grow because of faster productivity growth, perhaps due to greater entrepreneurship, as emphasized by our theory. They may also grow because of an expanding supply of housing, and Glaeser, Gyourko, and Saks (2006) found that human capital may predict increases in either the quantity or the price of housing, depending on the local regulatory environment. Finally, city growth responds to faster improvement in amenities, which skilled residents could induce through their demand as consumers and voters (Shapiro 2006). In Section 4, we use data on population growth, income growth and changes in housing values to estimate the extent of the power of these different forces. We find that the growth of skilled cities generally reflects growth in productivity rather than growth in amenities. The connection between growth and productivity seems strongest in the South and least strong in the West. The West is the only regions where skills are associated with increases in the quality of life. We also find that in the West, more skilled areas have had less housing supply growth, which may reflect that tendency of skilled people to organize to block new construction. We also try to separate out total productivity growth into growth in the number of employers and growth in the per-employer average productivity. We find that skills are more strongly correlated with growth in per-employer average productivity.

Section 5 turns to the connection between skills and urban resilience during the current recession. We look at the strong negative connection between skills and unemployment and find that this connection is larger than would be predicted solely on the basis of the cross-sectional relationship between education levels and unemployment rates. This fact is additional evidence for human capital spillovers at the city level, which may reflect the entrepreneurial tendencies of the more skilled. Section 6 concludes.

2 Ten Stylized Facts about Regional Decline and Resilience

We begin this paper with a broad perspective on urban resilience and change in the older areas of the United States. Our approach is non-standard. We follow economic historians such as Kim and Margo (2004) and take a very long perspective, going back, in some cases, to 1790. This longer perspective then forces us to focus on counties rather than cities or metropolitan areas. County data is available for long time periods, and while it is possible to use modern metropolitan definitions to group those counties, we believe that such grouping introduces a considerable bias into our calculations. Since metropolitan area definitions are essentially modern, we would be using an outcome to define our sample, which introduces bias. Low-population areas in the 19th century would inevitably have to grow unusually quickly if they were to be populous enough to be counted as metropolitan areas in the 20th century.

We also include only counties in the eastern and central portions of the United States, to avoid having our results dominated by the continuing westward tilt of the U.S. population. The western limit of our data is 90th meridian (west), the location of Memphis, Tennessee: Mississippi can be thought of as the data's western border. We also exclude those areas that are south of 30th parallel, which exclude much of Florida and two counties in Louisiana, and those areas north of the 43rd parallel, which exclude some northern areas of New England and the Midwest. While we will present data going back to the 1790 Census, we think of this area as essentially the settled part of the United States at the start of the Civil War, which allows us to treat the post-1860 patterns as essentially reflecting changes within a settled area of territory.

In this section, we examine ten stylized facts about regional change using this sample of counties. These facts inform our later theoretical discussion and may be helpful in other discussions of urban change. In some cases, these facts are quite similar to facts established using cities and metropolitan areas, but in other cases the county-level data display their own idiosyncrasies.

Fact # 1: Population patterns have been remarkably persistent over long time periods

Perhaps the most striking fact about this sample of counties is the similarity of population patterns in 1860 and today. When we regress the logarithm of population in 2000 on the logarithm of population in 1860, we find:

$$\log(\text{Pop in 2000}) = \underset{(.32)}{1.268} + \underset{(.03)}{.996} \cdot \log(\text{Pop in 1860}). \tag{1}$$

There are 1124 observations and the R^2 is .439, which corresponds to a 66 percent correlation. Population in 2000 rises essentially one-for-one with population in 1860, as shown in Figure 1. Some persistence is naturally to be expected because the housing stock of a city is durable (Glaeser and Gyourko, 2005). But our finding implies furthermore that, over this long time horizon, Gibrat's law operates: the change in population is essentially unrelated to the initial population level.

If we restrict ourselves to land even further east, using the 80th parallel as the boundary (about Erie, Pennsylvania), we estimate:

$$\log(\text{Pop in 2000}) = \underset{(.58)}{-.38} + \underset{(.06)}{.1.17} \cdot \log(\text{Pop in 1860}). \tag{1$'$}$$

In this case, there are only 306 observations, and the R^2 rises to .57, which represents a 75 percent correlation between population in 1860 and population in 2000 in this easternmost part of the U.S. While urban dynamics in America often seem quite volatile, there is a great deal of permanence in this older region. In this sample, there is a positive correlation between initial population levels and the rate of subsequent population growth, suggesting a tendency towards increased concentration.

Fact # 2: Population growth persists over short periods but not long periods

The permanence of population levels is accompanied by a remarkable permanence of population growth rates over shorter time periods. The first two columns of Table 1 show the correlation of population growth rates, measured with the change in the logarithm of population, and the lagged value of that variable. The first column shows results for our entire sample. The second column shows results when we restrict the sample to include only those counties that have 50,000 people at the start of the lagged decade.

Column 1 shows that in every decade, except for the 1870s, there is a strong positive correlation between current and lagged growth rates. Between the 1800s and the 1860s, the correlation coefficients range from .32 to .47. Then during in the aftermath of the Civil War there is a reversal, but starting in the 1880s, the pattern resumes again: between the 1880s and the 1940s, the correlation coefficients lie between .30 (the Great Depression decade) and .50 (the 1910s). During the post-war period, the correlations have been even higher, with correlation coefficients above .64 in all decades except for the somewhat unusual 1970s.

The pattern of persistence for more populous counties is even stronger. Over the entire period, the correlation coefficient never drops below .43. Except for the 1950s, the correlation coefficient is always higher for more populous counties than for smaller ones. The autocorrelation of growth rates for more populated counties was particularly high during the decades before the Civil War, when big cities were expanding rapidly in a more or less parallel path, and during more recent decades.

While short-term persistence is very much the norm for population growth rates, over longer periods growth rates can be negatively correlated. For the 54 counties that began with more than 50,000 people in 1860, an extra ten percent growth between 1860 and 1930 was associated with a lower 2.5 percent growth rate between 1930 and 2000, as shown by Figure 2. This negative correlation does not exist for the larger sample, but given that the persistence of decadal growth rates was even stronger among the counties with greater population levels the reversal is all the more striking. This negative relationship is our first indication of the changes in growth patterns over the 1860-2000 period. It suggests that different counties were growing during different epochs, and perhaps that fundamentally different forces were at work. We now turn to the relationship between initial population and later population growth, which is commonly called Gibrat's law.

Fact # 3: Gibrat's law is often broken

In studies of the post-war growth of cities and metropolitan areas, population growth has typically been found to be essentially uncorrelated with initial population levels both in the U.S. and elsewhere (Glaeser, Scheinkman and Shleifer 1995; Eaton and Eckstein 1997; Glaeser and Shapiro 2003). Gabaix (1999), Eeckhout (2004), and Córdoba (2008) have used this regularity to explain the size distribution of cities. Our long-run population persistence fact has already shown that Gibrat's law also seems to hold in our sample over sufficiently long time periods. In our entire sample, the correlation between change in log population between 1860 and 2000 is -.0034 and the estimated coefficient in a regression where change in the logarithm of population is regressed on the initial logarithm of population is -.0038 with a standard error of .033. There is also no correlation between the logarithm of population in 1950 and population change over the 50 years since then.

But Gibrat's law doesn't hold for many decades within our sample. Column 3 of Table

1 shows the correlation between the initial logarithm of population and the subsequent change in the logarithm of population over the subsequent decade. Column 4 shows the correlation only for more populous counties, those with at least 50,000 people at the start of the decade. The table shows that Gibrat's law holds during some time periods, but certainly not uniformly.

During the 19th century (with the exception of the 1860s), population growth is strongly negatively associated with initial population levels, especially in places that began with less people. This period is not marked by Gibrat's law at all—it is marked by mean reversion, as Americans spread out towards less populated counties. This process reflects improvements in transportation over this time period, and the great demand for newly accessible agricultural land.

While the entire sample is showing strong mean reversion, during the same period there is a positive, but usually insignificant, correlation between initial population levels and later growth in more populous counties. The pattern in this period is perhaps best understood as two separate processes that are going on simultaneously. Cities are getting bigger, as America grows, but empty farm areas are also gaining population.

This early period reflects the settlement of the region, and it can be considered anomalous and unrelated to patterns that should be expected to hold in a more mature area. We therefore focus more on the 20th century, when the eastern U.S. is more mature; but even then, Gibrat's law often fails to hold.

From the 1910s through 1960s, there was a long period where Gibrat's law, more or less, applies for more populated counties, but the larger sample shows faster population growth in places with higher initial levels of population. The process of centralized big city growth had become far weaker, but there was more growth in middle-population counties. This also reflects the relative decline of agriculture during those years and the fact that agriculture was overrepresented in the least dense counties.

Finally, from 1970 to 2000, the correlations between initial population and later growth are generally negative, especially in the most populous counties. This presumably reflects some of the impact of sprawl and the role that the automobile played in dispersing the American population.

Fact # 4: The 19th century moved west; the 20th century moved east

Just as Gibrat's law is hardly universal, there is also no universal pattern of horizontal movement within the region we consider. During the 19th century, the norm was to move west, but that reversed itself during much of the 20th century, within our restricted sample of counties. We focused on the eastern, central parts of the United States to reduce the impact of the enormous changes associated with the move to California and later to Florida. But that doesn't mean that there wasn't a westward push during much of 19th century. Table 2 shows the correlation between longitude and population growth by decade across our sample. Over the entire time period, there is no statistically significant correlation between population growth and longitude.

During every decade in the 19th century, growth was faster in the more western counties in our sample. This connection is strongest before the Civil War, when America is moving towards the Mississippi, but even as late as the 1890s, there is a weak negative relationship between longitude and population growth. To us, the more interesting fact is that since 1900, there is a move back east, at least in this sample. In every decade, except for the 1930s, longitude positively predicts growth. One interpretation of this fact was that the gains from populating the Midwest declined substantially after 1900, perhaps because America had become a less agricultural nation. According to this hypothesis, the eastern counties grew more quickly because they were better connected with each other and more suitable for services and manufacturing, and the agricultural communities declined. Since the 1970s, the connection between population growth and longitude has essentially disappeared.

Fact #5: The Great Lakes region grew during two distinct periods

In the early 19th century, waterways were the lifeline of America's transportation network, and the Great Lakes were the key arteries for the network. We calculate the distance between the county center and the center of the nearest Great Lake.[2] We then define proximity to the Great Lakes as the maximum of 200 minus the distance to the Great Lake centroid or zero.[3]

The second column of Table 2 shows the correlation between population growth and this measure of proximity to these large central bodies of water. Between 1790 and 1870, the correlation is uniformly positive, ranging from .07 during the 1850s to .44 during the 1810s. The early 19th century was the period when the Great Lakes had the strongest impact on population growth, which is not surprising since there were few other workable forms of internal transportation in the pre-rail era

Between 1870 and 1910, the correlation between proximity to the Great Lakes and growth is generally negative and quite weak. It turns out that this negative correlation is explained by the positive relationship between proximity to the Great Lakes and population levels in 1870 (.28 correlation coefficient). When we control for population in 1870, there is no

[2] We use ESRI Data & Maps 9.3 for the calculation.

[3] Unadjusted distance to the Great Lakes is strongly correlated with latitude and with temperature (-.89). Using a truncated measure, we can better distinguish proximity to the Great Lakes from coldness.

negative correlation between proximity to the Great Lakes and population growth between 1870 and 1900. Still, the absence of a positive relationship can be seen as an indication that the growing rail network had made access to waterways far less critical during the latter years of the 19th century.

Between 1910 and 1960, there is again a positive correlation between proximity to the Great Lakes and growth. Figure 3 shows the .33 correlation for counties within 200 miles of the Great Lakes. During this era of industrial growth and declining agricultural populations, factories grew in cities, like Detroit, that had once been centers of water-borne commerce. In some cases, the waterways were still important conduits for inputs and outputs. In other cases, industry located along the Great Lakes because this is where population masses were already located—about 44 percent of the positive correlation disappears when we control for population in 1910.

After this second surge of Great Lakes population growth, the region declined after 1970. Many explanations have been given for the decline of the Rust Belt, such as high union wages and an anti-business political environment (Holmes 1998), a lack of innovation in places with large plants and little industrial diversity, and the increasing desire to locate in sunnier climates. Our model below formalizes how technological progress has reduced the importance of logistical advantage conferred by the Great Lakes, thereby inducing a population shift towards regions with greater consumption amenities (Glaeser, Kolko, and Saiz 2001).

Fact # 6: The Sun Belt rose both after 1870 and after 1970

The third column in Table 2 shows the correlation between population growth and January temperature between 1790 and today. In every decade from the 1790s to the 1860s colder places show faster growth, as the North was gaining population relative to the South. Several factors explain this phenomenon. Many Northern areas had better farmland and they had a denser network of waterways. Industrialization came first to the North. Some illnesses, like malaria, were more prevalent in the South. For every extra degree of January temperature, population growth fell by .038 log points between 1810 and 1860, and by 1860, the correlation between county population levels and January temperature was -.41.

After the Civil War, the relationship between temperature and population growth reversed itself. In every decade from the 1870s to the 1900s, population growth was positively associated with January temperature. Every extra degree of January temperature was associated with .01 log points of growth between 1880 and 1910.[4] The effect of January temperature is strongest in less dense areas and the effect disappears in more populous counties.

[4]The 1870 Census is potentially problematic because of an undercount in the South (Farley 2008).

This may reflect higher fertility among the poorer and less educated Southern population (Steckel 1978). Increasing rail densities in the South may also have made farming in more remote areas more attractive.

The relationship between January temperature and population growth then disappears between 1910 and 1970. The correlation is weak, and if anything negative. Moreover, it is largely explained by the positive correlation between initial population levels and later growth: controlling for initial population, the effect of January temperature on growth during the entire 1910-1970 time period is indistinguishable from zero. The coefficients become significantly positive once we restrict the sample to counties with more than 50,000 people in 1910, consistent with previous evidence on city growth (Glaeser and Tobio 2008). Before 1970, people were moving to warmer cities, but not to warmer rural areas.

The three decades since have seen a remarkable rise of the Sun Belt. From 1970 to 2000 warmth is a strong positive predictor of population growth for all counties, and ten extra degrees of January temperature are associated with an extra .1 log points of population growth

In Table 3, we show the impact of initial population, January temperature, proximity to the Great Lakes and longitude in multivariate regressions for six different thirty-year periods.[5] Differences across columns remind us that all variables had different impacts in different epochs, and that regional growth can only be understood by bringing in outside information about changing features of the U.S. economy.

In the antebellum era, U.S. population was spreading out: proximity to the Great Lakes had a positive impact on growth, while longitude, January temperature and especially initial population had a negative impact. The overall explanatory power of these variables drops significantly for the late 19th century. Warmer areas grew more quickly, although the undercounting of Southern population in the 1870 Census means that this coefficient should be cautiously interpreted. January temperature also had a positive effect on population growth from 1900 to1930, but so did proximity to the East Coast and to the Great Lakes. Places with more initial population grew more quickly, reflecting the growth of big cities during those decades. Results for 1940-1970 are quite similar, except that January temperature is no longer significant. After 1970, January temperature becomes the most powerful predictor of county-level growth. Population moves east rather than west. Initial population is negatively associated with growth, which presumably reflects the growth of sprawl. Proximity to the Great Lakes has a slight negative impact on county-level population growth.

For the post-war period we also have income data that can help us make more sense of

[5] We skip the 1860s, which are unusual because of the Civil War, and the 1930s, which are unusual because of the Great Depression.

the growth of the South during this time period. Table 4 shows the correlation of county median incomes and other variables.[6] The correlation between income growth and January temperature is highest in the 1950s and 1960s, when the connection between January temperature and population growth is weakest. During this era, the Sun Belt was getting much more prosperous but it wasn't attracting a disproportionate number of migrants. After 1970, the connection between January temperature and income drops considerably, though the correlation between population growth and January temperature rises. One explanation for this phenomenon, given by Glaeser and Tobio (2008) is that over the last 30 years, sunshine and housing supply have gone together. The South seems to be considerably more permissive towards new construction, which may well explain why three of the fastest growing American metropolitan areas since 2000 are in states of the old Confederacy (Atlanta, Dallas and Houston).

Fact # 7: Income mean reverts

One explanation for Gibrat's law is that areas receive productivity shocks that are proportional to current productivity (Eeckhout 2004). But that interpretation is difficult to square with the well-known convergence of regional income levels found by Barro and Sala-i-Martin (1991) and others. In our data sample, median incomes also mean revert. We have data on median income levels starting in 1950, and the second column of Table 4 shows the correlation between the decadal change in the logarithm of this variable and the logarithm of the variable.

The table shows that during every decade except the 1980s, income growth was substantially lower in places that started with higher income levels. Overall, if median income was .1 log points higher in 1950, it grew by .066 log points less from 1950 to 2000, as Figure 4 shows. Income in 1950 can explain 72 percent of the variation in income since then. While population levels persist, income levels generally do not.

Income convergence does seem to have fallen off after 1980, most notably during the 1980s and among larger cities. In the whole sample, as income in 1980 rises by .1 log points income growth from 1980 to 2000 falls by .0049 log points. But the relationship is instead positive for counties that began with more than 50,000 people. This weakening of income convergence may be due to an increase in the returns to skill.

There is a positive correlation between population growth and initial income levels which may explain some of the income convergence. Between 1950 and 1980, an extra .1 log points of initial income was associated with a reduction in income growth of .06 log points and an increase in population growth of .03 log points. But given conventional estimates of labor

[6] This income measure does nothing to control for the human capital composition of the population.

demand elasticity (Borjas 2003), this population growth can only explain about a fifth of income convergence. Other explanations for income convergence are that technology has spread over space, and capital mobility and changing composition of the labor force. The last explanation, however, is troubled by the fact that the share of the population with college degrees has increased more quickly in places that had higher incomes in 1950; on average a .1 log point increase in 1950 incomes is associated with a .007 percent increase in the share of the adult population with college degrees.

Since 1980, higher income growth in one decade predicts lower income growth over the next ten years. This can be reconciled with the strong positive persistence of population growth if a steady flow of new people is pushing wages down in some areas.

Fact # 8: Manufacturing predicts the decline of cities but not the decline of counties

Many papers have noted the negative correlation between concentration in manufacturing and subsequent urban growth (Glaeser, Scheinkman and Shleifer 1995). This correlation does not appear in our county data. We use the share of the county's employment that is in manufacturing in 1950 as our measure of the concentration of the county in manufacturing at the start of the post-war era.

Figure 5 shows that as the share of a county's workers in manufacturing in 1950 rises by 10 percent, subsequent population growth rises by .07 log points. The effect grows stronger if we control for initial population, January temperature and proximity to the Great Lakes. It gets slightly weaker if we control for initial income, because manufacturing counties had higher wages.

This positive correlation does not hold for the more populous counties, which presumably explains why city and metropolitan-area data show a negative connection between manufacturing and growth. If we restrict our sample to counties with more than 100,000 people in 1950, the correlation becomes negative. Manufacturing left cities, and cities that were highly concentrated in manufacturing declined.

At the county level, an initial concentration in manufacturing does not seem to have had such a negative impact. It does predict income decline in every decade except the 1980s: the last column in Table 4 shows that a 10 percent rise in the share of manufacturing in 1950 is associated with a .114 log point fall in median incomes between 1950 and 2000. However, once we control for initial income, manufacturing is positively associated with income growth as well as population growth.

Again, the impact of manufacturing on income growth is more strongly negative in more populous counties. This reinforces the view that manufacturing has proven to be far worse for densely populated areas than for those with fewer people. Big factories seem a better match for moderate density levels (Glaeser and Kohlhase 2004).

Fact # 9: Education predicts post-war growth

A series of papers have also shown the connection between education and the success of cities (Rauch 1993; Glaeser, Scheinkman and Shleifer 1995; Simon and Nardinelli 2002; Glaeser and Saiz 2004; Shapiro 2006). We now ask whether this correlation also holds at the county level. Table 5 shows the correlation between the share of the adult population with college degrees and subsequent income and population growth. We have this during every decade except 1960, and for that year, we use the college attainment rates in 1950 instead.

The first column shows that college attainment and population growth have a strong positive correlation in our sample. In the long run, as the share of the population with college degrees increases by 10 percent in 1940, population growth between 1940 and 2000 increases by .13 log points. Over shorter periods, the positive effect is strongest in the 1950s and 1960s, and it holds in every decade but the 1970s, when there is a negative correlation that becomes insignificant when we control for the logarithm of 1970 population.

The second column shows results for income growth. Across the entire sample, there is a negative relationship between initial education and subsequent income growth. This is certainly not true across cities or metropolitan areas. Across counties, the effect is primarily due to mean reversion in median incomes. Controlling for initial log income, the estimated coefficients for the initial share of the population with college degrees are always positive, and they are statistically significant for the 1950s (.89), the 1960s (.55, using college attainment in 1950), and the 1980s (.9). More educated places seem to be growing both in population and income, once we account for the tendency of incomes to revert to the mean.

Glaeser and Resseger (2010) present evidence suggesting that skills have more impact in larger cities. In theory, urban density is more valuable when it connects people who have more to teach one another. The last two columns of Table 5 focus on those counties that begin the decade with at least 100,000 people. Column 3 shows the population growth correlations, which are uniformly positive, but not always larger than those observed in the entire sample. Column 4 shows that the correlation between income growth and education is always more positive for more populous counties than for the entire sample. In the 1950s and 1960s, when skills were negatively associated with income growth in the entire sample, skills were positively associated with income growth in more populous counties. These results support the view that there is a complementarity between skills and density.

In Table 6, we present two regressions looking at the entire 1950-2000 period. In the first regression, income growth is the dependent variable. In the second regression, population growth is the dependent variable. We include as controls January temperature, longitude and distance to the Great Lakes. We control for the logarithms of initial education and population. We also include the share of employment in manufacturing, the share of the population with college degrees and an interaction between the logarithm of 1950 population and the share of the population with college degrees. We have normalized the initial population by subtracting the mean of that variable in this sample; this enables us to glean the impact of education for the mean city with the coefficient in the regression.

Initial income strongly predicts subsequent income declines and significant population increases. Initial population is negatively associated with both income and population growth. Proximity to the East Coast, longitude and manufacturing are both positively correlated with both income and population growth. Proximity to the Great Lakes has no impact on population growth, but a negative correlation with income growth.

Education has a positive effect on both income and population growth. At the average initial population level, as the share of adults with college degrees in 1950 increases by 3 percent (about one standard deviation), subsequent population growth increases by slightly more than .12 log points (about 12 percent) and income growth rises by around 7 percent. These effects are statistically significant and economically meaningful.

The effects of education on income and population growth are stronger for counties with higher initial levels of population. As the level of population increases by one log point (slightly less than one standard deviation), the impact of education on population growth increases by 54 percent and the impact of education on income growth increases by 36 percent. Skills do seem, over the fifty year period, to have had a particularly strong positive effect on income and population growth for areas that initially had higher levels of population.

While it is clear that skills matter during the post-war period, it is less clear whether skills were as important before World War II. We are limited by an absence of good education data during this earlier period, which is why Simon and Nardinelli (2002) focus on the presence of skilled occupations in 1900. Yet because it seems worthwhile to know whether skilled places also grew in the 19th century, Table 7 shows the correlation between the share of the population with college degrees in 1940 and growth over the entire 1790-2000 period. There are at least two major problems with this procedure. First, skill levels change, and a place that is skilled in 1940 may well not have been skilled in 1840. We are only moderately reassured by the .75 correlation between the share of the population with college degrees in 1940 and the share of the adult population with college degrees in 2000. Second, it is possible that skilled people came disproportionately to quickly growing areas. Indeed, there is a strong positive correlation (.61) in our sample between population growth between 1940 and 2000 and the growth in the share of the population with college degrees over the same time period.

Despite these caveats, Table 7 shows the correlations over the long time period. The first column includes all of our counties; the second column shows results only for those counties with more than 50,000 people at the start of the decade. The table shows a strong positive correlation between skills in 1940 and growth in population for most of the twentieth century. In the 19th century, education was largely uncorrelated with growth across the entire sample. Among more populous counties, the correlation is generally positive after 1820. One interpretation of these differences is that there was a complementarity between cities and skills even in the 19th century. A second interpretation is that skills in 1940 are a reasonable proxy for skills in the 19th century among more populous counties, but not for sparsely populated areas that presumably changed more over the century.

Those different interpretations yield different conclusions about the long run correlation between skills and population growth. If the latter interpretation is correct, and the correlation disappears because skills in 1940 don't correlate with 19th century skills, then the skills-growth correlation may be the one relationship that holds virtually over our entire sample. If, however, the former interpretation is correct, then the relationship between skills and growth is, like everything else we've looked at, a phenomenon that holds only during certain eras.

Moretti (2004) and Berry and Glaeser (2005) report a positive correlation between initial levels of education and education growth over the post-war period. We confirm this powerful fact with our cross-county data. We look at the relationship between change in the share of population with college degrees between 1940 and 2000 and the share of the population with college degrees in 1940. Over the entire sample, we estimate the relationship:

$$\text{Change in share with BAs 1940–2000} = \underset{(.003)}{.048} + \underset{(.088)}{2.66} \cdot \text{Share with BAs in 1940.} \tag{2}$$

Standard errors are in parentheses. There are 1326 observations and the R^2 is .4. As the share with college degrees in 1940 increases by 2 percent, growth in the share of college degrees increases by 5.32 percent. Figure 6 illustrates this relationship.[7] The only decade in which there is no positive correlation between initial schooling and subsequent growth in schooling is the 1940s. Afterwards, schooling uniformly predicts schooling growth. In the 1970s, 1980s, and 1990s, the correlation coefficients between initial schooling and subsequent increases in the share with college degrees are .57, .66 and .54 respectively. One of the reasons why initially skilled places have done so well is quite possibly that they have attracted more

[7] To make the graph less cluttered, we only display counties with at least 50,000 people in 1940.

skilled people over time.

Fact # 10: Firm size is strongly correlated with employment and income growth after 1980

Glaeser et al. (1992) found a strong negative correlation between average firm size and subsequent growth across large industrial groups within metropolitan areas. Glaeser, Kerr and Ponzetto (2010) show that smaller firm size predicts growth both across and within metropolitan areas. Our last fact is that firm size is correlated with population and income growth across our sample of counties.

Firm size is typically measured by looking at the ratio of the number of establishments to the number of employees within a metropolitan area or industrial cluster. In our case, we use the 1977 County Business Patterns data and calculate the average number of employees per establishment in each county in our sample. The variable ranges from 2.9 to 35, with a sample mean of 12.74. There is a strong positive correlation between county population and average establishment size.

Table 8 shows four growth regressions that include average establishment size. The first two look at population growth between 1980 and 2000. Columns 3 and 4 show results on growth in median income over the same two decades. Columns 1 and 3 look at our entire sample. Columns 2 and 4 look only at those counties that had at least 50,000 people in 1980. In all cases, we include our standing controls including the logarithms of initial income and population, the share of the labor force in manufacturing, our geographic controls and the initial share of the population with a college degree. The effect of these variables is unchanged from our previous regressions.

Regressions 1 and 2 both show the strong negative correlation between average establishment size and subsequent population growth. As average establishment size rises by four workers (approximately one standard deviation), subsequent population growth declines by .06 log points across the entire sample. The effect is somewhat larger for more populous counties, where the decline is around 10 percentage points.

Regressions 3 and 4 show the strong negative connection between average establishment size and income growth. As average establishment size increases by four, income growth declines by .045 log points across the entire sample, and by .06 log points in the sample of more populous counties. These effects are comparable in magnitude with the education effect on income growth and even stronger statistically.

While larger establishment sizes do seem to predict less growth of income and population, it is less clear how to interpret these facts. Glaeser et al. (1992) interpreted the positive connection between small firm size and later growth as evidence on the value of competition.

Miracky (1995) observed the same phenomenon and associated it with the product life cycle. While this remains one plausible interpretation, the fact that these connections occur within very finely detailed industry groups, and controlling for average establishment age, speaks against this interpretation. Glaeser, Kerr and Ponzetto (2010) suggest that these connections suggest the value of local entrepreneurship. We prefer this latter interpretation, which will fit closely with the following model, but we certainly acknowledge that other interpretations are possible. We also recognize that entrepreneurship has received multiple definitions and has proven difficult to measure empirically. Our ultimate focus is on entrepreneurs as the drivers of change, innovation and productivity growth (Audretsch 1995). In practice, such entrepreneurial activity has been commonly proxied by business ownership rates and by the creation of new firms, while small firms have been increasingly recognized as key contributors to innovation (Audretsch 2003).

In the last two columns of Table 7, we also look at the correlation between firm size and growth during early decades. We use average establishment size in 1977, an ex-post measure that raises all the concerns we had about using schooling in 1940 to proxy for education in the 19th century. In this case, the negative relationship between firm size in 1977 and growth is not present during earlier decades. Either the small firm size effect is specific to the past thirty years, or small firm size in 1977 doesn't capture small firm size during earlier years. Certainly, when Glaeser et al. (1992) looked at firm size in 1957, they found a negative correlation with subsequent growth.

3 Theoretical Framework

We now present a model of regional change, skills and resilience. The model provides a framework that will enable us to understand better the reasons why skilled areas have grown more quickly over the past sixty years. In principle, it is possible that skilled places could have been growing more quickly because of improvements in productivity, amenities or housing supply. We need a formal framework to help separate these competing explanations. The model will also deliver some intuition as to why skills have been so important in the older areas of the U.S. that seems to have been hit by adverse shocks after World War II.

Individual utility is defined over consumption of land, denoted L, and a CES aggregate of measure G of differentiated manufactured goods, each denoted $c(\nu)$. Thus

$$U = \theta_i \left[\int_0^G c(\nu)^{\frac{\sigma-1}{\sigma}} d\nu \right]^{\frac{\mu\sigma}{\sigma-1}} L^{1-\mu}, \tag{3}$$

where $\theta_i > 0$ is a quality of life multiplier associated with the exogenous amenities of city i.

Each entrepreneur manufactures a differentiated variety employing labor according to the linear production function

$$x(\nu) = a_i n(\nu), \tag{4}$$

where $x(\nu)$ is the output of firm ν, $n(\nu)$ its workforce, and a_i the productivity of entrepreneurs in city i. The Appendix derives the optimal pricing and hiring decisions of monopolistically competitive manufacturers.

City i is endowed with an exogenous number of entrepreneurs, denoted E_i. With an endogenous workforce of N_i full-time workers, its equilibrium wage is

$$w_i = \frac{\sigma - 1}{\sigma} \left(\mu Y a_i^{\sigma-1} \frac{E_i}{N_i} \right)^{\frac{1}{\sigma}}, \tag{5}$$

having normalized to unity the price index for the composite manufactured good.

City i has a fixed quantity of land, denoted by $\bar{L}_i$, which is owned by developers who reside in the city itself. Given the utility function (3), workers, entrepreneurs, and developers all choose to spend a fraction $1 - \mu$ of their income on consumption of land. Hence equilibrium in the real-estate market implies that the price of land in city i is

$$r_i = \frac{1 - \mu}{\mu \bar{L}_i} \left[\mu Y E_i (a_i N_i)^{\sigma-1} \right]^{\frac{1}{\sigma}}. \tag{6}$$

In an open-city model in which workers are fully mobile, their utility needs to be equalized across locations. Spatial equilibrium then requires

$$\theta_i w_i r_i^{\mu-1} = \theta_j w_j r_j^{\mu-1} \text{ for all } i, j. \tag{7}$$

We consider a continuum of cities, each of which is arbitrarily small compared to the aggregate economy. Then, letting $N = \int N_j dj$ denote the aggregate size of the workforce, for each city i the equilibrium workforce is

$$\log N_i = \kappa_N + \frac{\sigma \log \theta_i + \mu(\sigma - 1) \log a_i + \mu \log E_i + (1 - \mu) \sigma \log \bar{L}_i}{\mu + \sigma - \mu\sigma}, \tag{8}$$

where the constant κ_N is independent of idiosyncratic shocks affecting the city. Likewise, for given constants κ_w and κ_r equilibrium wages are

$$\log w_i = \kappa_w + \frac{(1 - \mu)\left[(\sigma - 1) \log a_i + \log E_i - \log \bar{L}_i\right] - \log \theta_i}{\mu + \sigma - \mu\sigma}, \tag{9}$$

and equilibrium rents

$$\log r_i = \kappa_r + \frac{(\sigma - 1)(\log \theta_i + \log a_i) + \log E_i - \log \bar{L}_i}{\mu + \sigma - \mu\sigma}. \tag{10}$$

The three constants κ_N, κ_w and κ_r are defined exactly in the Appendix

A disaggregation of city-specific productivity into separate components for manufacturing and logistical efficiency enables the model to provide a simple account of the role of transport costs in the pattern of U.S. regional dynamics during the latter part of the twentieth century. Specifically, let productivity in city i at time t be

$$a_{i,t} = A_{i,t} \exp\left(-\frac{\Gamma_i}{T_t}\right). \tag{11}$$

In this decomposition, $A_{i,t}$ captures the productive efficiency achieved at time t by entrepreneurs in city i, measured by output per worker in their firms. However, delivering goods to the final consumer involves transportation and distribution costs such that for every x units shipped from a plant in city i, only $x \exp(-\Gamma_i/T_t)$ reach the final consumer, according to the conventional specification of "iceberg" transport costs. The time-invariant city-specific parameter $\Gamma_i > 0$ is a measure of each city's natural logistical advantages, resulting from geographic characteristics such as access to waterways. The time-varying common parameter T_t measures the ability of transportation technology to overcome natural obstacles. The following result then obtains.

Proposition 1 *Advances in transportation technology reduce the share of the cross-city variance of population, income, and housing prices that is explained by heterogeneity in natural logistical advantages:*

$$\frac{\partial^2 Var(\log \Xi_{i,t})}{\partial T_t \partial Var(\Gamma_i)} < 0 = \frac{\partial^2 Var(\log \Xi_{i,t})}{\partial T_t \partial Var(\theta_{i,t})} = \frac{\partial^2 Var(\log \Xi_{i,t})}{\partial T_t \partial Var(A_{i,t})} = \frac{\partial^2 Var(\log \Xi_{i,t})}{\partial T_t \partial Var(E_{i,t})} = \frac{\partial^2 Var(\log \Xi_{i,t})}{\partial T_t \partial Var\left(\bar{L}_{i,t}\right)}$$

for all $\Xi_{i,t} \in \{N_{i,t}, w_{i,t}, r_{i,t}\}$.

Our stylized fact 5 emphasized that the rise of the Great Lakes region was due to the crucial importance of proximity to the waterways through which most domestic trade used to flow. Over time, technological progress was a substitute for a favorable location: as transportation technology improved (an increase in T_t), natural harbors and geographic accessibility (Γ_i) came to matter less for regional success. Heterogeneity in amenities ($\theta_{i,t}$), housing supply ($\bar{L}_{i,t}$) and entrepreneurial achievement ($A_{i,t}$ and $E_{i,t}$) then acquired proportionally greater importance, leading to the rise of the Sun Belt (fact 6) and the enduring success of cities and regions with the highest levels of human capital (fact 9).

Through equations (8), (9), and (10), the model provides us with the basis for our empirical work in Section 4. We assume that for each city i and time t the values of $\theta_{i,t}$, $a_{i,t}$, $E_{i,t}$ and $\bar{L}_{i,t}$ evolve according to the dynamics

$$\theta_{i,t+k} = \theta_{i,t} \exp\left(k\boldsymbol{\beta}^{\theta} \cdot \mathbf{X}^{i} + \varepsilon^{\theta}_{i,t+k}\right), \tag{12}$$

$$a_{i,t+k} = a_{i,t} \exp\left(k\boldsymbol{\beta}^{a} \cdot \mathbf{X}^{i} + \varepsilon^{\tau}_{i,t+k}\right), \tag{13}$$

$$E_{i,t+k} = E_{i,t} \exp\left(k\boldsymbol{\beta}^{E} \cdot \mathbf{X}^{i} + \varepsilon^{E}_{i,t+k}\right) \tag{14}$$

and

$$\bar{L}_{i,t+k} = \bar{L}_{i,t} \exp\left(k\boldsymbol{\beta}^{\bar{L}} \cdot \mathbf{X}^{i} + \varepsilon^{\bar{L}}_{i,t+k}\right). \tag{15}$$

The parameter vectors $\boldsymbol{\beta}^{\theta}$, $\boldsymbol{\beta}^{a}$, $\boldsymbol{\beta}^{E}$ and $\boldsymbol{\beta}^{\bar{L}}$ connect time-invariant city characteristics, denoted by $\mathbf{X}^{i}$, with growth in θ, a, E and $\bar{L}$ respectively. The terms $\varepsilon^{\theta}_{i,t+k}$, $\varepsilon^{a}_{i,t+k}$ $\varepsilon^{E}_{i,t+k}$, and $\varepsilon^{\bar{L}}_{i,t+k}$ are stochastic errors.

For any set of variables $\mathbf{X}^{i}$ we can then write

$$\log N_{i,t+1} - \log N_{i,t} = \frac{\sigma\boldsymbol{\beta}^{\theta} + \mu\left(\sigma-1\right)\boldsymbol{\beta}^{a} + \mu\boldsymbol{\beta}^{E} + \left(1-\mu\right)\sigma\boldsymbol{\beta}^{\bar{L}}}{\mu+\sigma-\mu\sigma} \cdot \mathbf{X}^{i} + \varepsilon^{N}_{i,t}, \tag{16}$$

and

$$\log w_{i,t+1} - \log w_{i,t} = \frac{\left(1-\mu\right)\left[\left(\sigma-1\right)\boldsymbol{\beta}^{a} + \boldsymbol{\beta}^{E} - \boldsymbol{\beta}^{\bar{L}}\right] - \boldsymbol{\beta}^{\theta}}{\mu+\sigma-\mu\sigma} \cdot \mathbf{X}^{i} + \varepsilon^{w}_{i,t}, \tag{17}$$

where $N_{i,t}$ and $w_{i,t}$ are the number of workers and the wage level in city i at time t, and $\varepsilon^{N}_{i,t}$ and $\varepsilon^{w}_{i,t}$ are error terms.

We could perform a similar first difference for housing costs, but our data on real estate typically involve home prices, which are a stock of value rather than a flow. The stock value of land in our model at time t, denoted $V_{i,t}$, can be interpreted as the discounted value of the flow of future land rents or future flow costs:

$$V_{i,t} = \mathbb{E}\left(\int_{k=0}^{\infty} e^{-\rho k} r_{i,t+k} dk\right) = r_{i,t}\mathbb{E}\left(\int_{k=0}^{\infty} e^{(g_r-\rho)k+\varepsilon^{r}_{t+k}} dk\right), \tag{18}$$

where

$$g_r \equiv \frac{\left(\sigma-1\right)\left(\boldsymbol{\beta}^{\theta} + \boldsymbol{\beta}^{a}\right) + \boldsymbol{\beta}^{E} - \boldsymbol{\beta}^{\bar{L}}}{\mu+\sigma-\mu\sigma} \cdot \mathbf{X}^{i} \tag{19}$$

is the time-invariant expected growth rate of future rents and ε^{r}_{t+k} the relative error term.

For a time-invariant error distribution,

$$\log V_{i,t+1} - \log V_{i,t} = \log r_{i,t+1} - \log r_{i,t} = \frac{(\sigma - 1)\left(\boldsymbol{\beta}^{\theta} + \boldsymbol{\beta}^{a}\right) + \boldsymbol{\beta}^{E} - \boldsymbol{\beta}^{\bar{L}}}{\mu + \sigma - \mu\sigma} \cdot \mathbf{X}^{i} + \varepsilon_{i,t}^{V}. \quad (20)$$

If then, we have estimated coefficients for a variable, such as schooling, in population, income, and housing-value growth regressions of B_{Pop}, B_{Inc}, and B_{Val} respectively, then by combining these estimated coefficients it is possible to uncover the underlying connections between a variable and growth in amenities, land availability, and entrepreneurship. Algebra yields the effect on residential amenities

$$\beta_s^{\theta} = -B_{Inc} + (1 - \mu)\, B_{Val}, \quad (21)$$

on the supply of real estate

$$\beta_s^{\bar{L}} = B_{Pop} + B_{Inc} - B_{Val}, \quad (22)$$

and on productivity-increasing entrepreneurship

$$\tilde{\beta}_s^{E} \equiv (\sigma - 1)\, \beta_s^{a} + \beta_s^{E} = B_{Pop} + \sigma B_{Inc}. \quad (23)$$

The last coefficient captures both the extensive margin of entrepreneurship, which corresponds to the creation of more numerous firms, and its intensive margin, namely the creation of more efficient firms. The two components can be disentangled through their different impact on average firm size, measured by employment per firm $n_i = N_i/E_i$, which evolves as

$$\log n_{i,t+1} - \log n_{i,t} = \frac{\sigma\boldsymbol{\beta}^{\theta} + \mu(\sigma - 1)\,\boldsymbol{\beta}^{\tau} + (1 - \mu)\,\sigma\left(\boldsymbol{\beta}^{\bar{L}} - \boldsymbol{\beta}^{E}\right)}{\mu + \sigma - \mu\sigma} \cdot \mathbf{X}^{i} + \varepsilon_{i,t}^{n}. \quad (24)$$

If an additional average firm size growth regression yields an estimated coefficient of B_{Siz}, we can infer that the effect on the intensive margin is

$$\beta_s^{a} = \frac{\sigma B_{Inc} + B_{Siz}}{\sigma - 1}, \quad (25)$$

and the effect on the extensive margin is

$$\beta_s^{E} = B_{Pop} - B_{Siz}. \quad (26)$$

3.1 Endogenous Entrepreneurship and Responses to Shocks

While the previous equations will serve to frame our empirical work in Section 4, we now focus on the connection between skills, entrepreneurship and regional resilience. An adverse regional shock can be understood as a reduction Δ_i in the exogenous stock of entrepreneurs $\bar{E}_i$, due to death or technological obsolescence or migration, so only $\bar{E}_i - \Delta_i$ entrepreneurs remain. The ability of a region to respond to such a shock will depend on the production of new ideas. To address this, we endogenize entrepreneurship, and assume that all workers are endowed with one unit of time that they can spend either working or engaging in entrepreneurial activity. The time cost of trying to become an entrepreneur is a fixed quantity t. If the worker becomes an entrepreneur, she has an individual-specific probability η of being successful. The value of an entrepreneurial attempt is thus $\eta\pi_i + (1-t)\,w_i$.

We assume that there is a distribution of η in the population such that the share of agents with probability of success no greater than η equals η^α for $\alpha \in (0,1)$.[8] Given this assumption, suppose that city i has a number M_i of potential entrepreneurs. All those with probabilities of success greater than $\bar{\eta}_i$ attempt entrepreneurship, while those with probability of success below $\bar{\eta}_i$ spend all their time as employees. Then the total number of entrepreneurs equals

$$E_i = \bar{E}_i - \Delta_i + \frac{\alpha}{1+\alpha}\left(1-\bar{\eta}_i^{1+\alpha}\right)M_i, \tag{27}$$

while the labor supply is

$$N_i = \left[1 - t\left(1-\bar{\eta}_i^{\alpha}\right)\right]M_i. \tag{28}$$

These in turn determine wages w_i and firm profits π_i, as detailed in the appendix. It is privately optimal for an agent to attempt entrepreneurship if and only if his probability of success is $\eta\pi_i \geq tw_i$. Thus an equilibrium is given by

$$\bar{\eta}_i = 1 \text{ if } M_i \leq (\sigma-1)\,t\left(\bar{E}_i - \Delta_i\right), \tag{29}$$

and if instead $M_i > (\sigma-1)\,t\left(\bar{E}_i - \Delta_i\right)$, by

$$\bar{\eta}_i \in [0,1] \text{ such that } \bar{\eta}_i = \frac{(\sigma-1)\,t}{1-t\left(1-\bar{\eta}_i^{\alpha}\right)}\left[\frac{\bar{E}_i - \Delta_i}{M_i} + \frac{\alpha}{1+\alpha}\left(1-\bar{\eta}_i^{1+\alpha}\right)\right]. \tag{30}$$

which is uniquely defined since the right-hand side is a monotone decreasing function of $\bar{\eta}_i$.

In particular if $t = 1$, so people are either would-be entrepreneurs or employees, then the following result holds for a closed city with an exogenous number $\bar{M}_i$ of agents choosing between employment and entrepreneurship.

[8] In other words, $1/\eta$ has a Pareto distribution with a minimum of 1 and shape parameter α.

Proposition 2 *In a closed city, both wages and the number of employers fall in response to a negative shock* ($\partial \log w_i / \partial \Delta_i < 0$ *and* $\partial \log E_i / \partial \Delta_i < 0$)*, but their proportional decline is smaller in magnitude if the endogenous supply of entrepreneurs is more elastic* ($\partial^2 \log w_i / (\partial \alpha \partial \Delta_i) \geq 0$ *and* $\partial^2 \log E_i / (\partial \alpha \partial \Delta_i) \geq 0$)*.*

Proposition 2 delivers the connection between urban resilience and entrepreneurship in a closed-city framework. As older employers either go bankrupt or leave the city, this causes incomes in the city to decline. This negative shock can be offset by entrepreneurship, as a decline in wages causes entrepreneurship to become relatively more attractive. If the supply of entrepreneurship is more elastic, which is captured by a higher value of the parameter α, then there is a stronger entrepreneurial response to urban decline and the impact of a negative shock on incomes becomes less severe.

The closed-city model also allows us to shed light on the observed correlation between urban resilience and cross-city differences in average firm size.

Proposition 3 *Consider a set of closed cities with identical size,* $M_i = \bar{M}$ *for all* i*. Both wages and the number of employers fall in response to a negative shock* ($\partial \log w_i / \partial \Delta_i < 0$ *and* $\partial \log E_i / \partial \Delta_i < 0$)*, but their proportional decline is smaller in magnitude in cities with a lower initial average firm size* ($\partial^2 \log w_i / (\partial \Delta_i \partial n_i) < 0$ *and* $\partial^2 \log E_i / (\partial \Delta_i \partial n_i) < 0$)*.*

Keeping city size constant, higher firm density and smaller average firm size are the indication of greater entrepreneurship. When a negative shock hits, some firms are forced to shut down by exogenous forces such as the obsolescence of their product or the death of an entrepreneur. Although cushioned by the entry of new entrepreneurs, this blow implies a fall in the number of employers and in the local wage level. Intuitively, the crisis is more severe in cities that did not have a diversified set of firms to begin with, because those cities are reliant on a few large employers and thus suffer disproportionately from the disappearance of any single firm.

To extend our analysis to the open-city model, we assume that $t = 0$, so there is no time cost to entrepreneurship. In this case, everyone tries to be an entrepreneur, which means that $\bar{\eta} = 0$. In a closed city, it would remain true that $\partial w_i / \partial \Delta_i < 0$ and $\partial^2 \log w_i / (\partial \alpha \partial \Delta_i) > 0$, so a greater endogenous supply of entrepreneurs offsets the negative effects of an exogenous shock to the number of employers. When the city is open, we assume that people choose their location before the realization of their individual entrepreneurial ability η. Spatial equilibrium then requires $\theta_i y_i r_i^{\mu - 1} = \bar{U}$ for all i, where $y_i \equiv w_i + \pi_i \alpha / (1 + \alpha)$ denotes expected earnings. With a continuum of atomistic cities, the following result holds.

Proposition 4 *Expected earnings, the total number of employers, and the price of land decrease in the exogenous negative shock to the endowment of employers* ($\partial y_i/\partial\Delta_i < 0$, $\partial E_i/\partial\Delta_i < 0$, *and* $\partial r_i/\partial\Delta_i < 0$) *and increase in the endogenous supply of entrepreneurs* ($\partial y_i/\partial\alpha > 0$, $\partial E_i/\partial\alpha > 0$, *and* $\partial r_i/\partial\alpha > 0$). *The labor supply and city population* ($\Lambda_i \equiv \bar{E}_i - \Delta_i + N_i$) *increase in the endogenous rate of entrepreneurship* ($\partial\Lambda_i/\partial\alpha = \partial N_i/\partial\alpha > 0$). *If the endogenous supply of entrepreneurs is sufficiently elastic, population decreases with an exogenous negative shock to the endowment of employers* ($\alpha \geq 1/(\sigma - 1) \Rightarrow \partial\Lambda_i/\partial\Delta_i < 0$).

In the limit case $\mu = 1$, *the labor supply and city population both decrease with an exogenous negative shock to the endowment of employers* ($\partial\Lambda_i/\partial\Delta_i < \partial N_i/\partial\Delta_i < 0$). *Moreover, a greater endogenous supply of entrepreneurship mutes the proportional impact of a negative endowment shock on expected earnings, the total number of employers and city population* ($d^2 \log y_i/(d\alpha d\Delta_i) > 0$,$d^2 \log E_i/(d\alpha d\Delta_i) > 0$, *and* $d^2 \log \Lambda_i/(d\alpha d\Delta_i) > 0$).

Proposition 4 makes the point that entrepreneurship can substitute for a decline in an area's core industries in a way that keeps population, earnings, and real-estate values up. A higher rate of exodus for older industries will cause a city to lose both population and income, but that can be offset if the city also has a higher rate of new entrepreneurship.

What factors are likely to make entrepreneurship more common? One possibility is skilled workers have a comparative advantage at producing new ideas. To capture this intuition, we assume that there are two types of workers. Less skilled workers have one unit of human capital and have a value of $\alpha/(1+\alpha)$ equal to $\underline{\alpha}$. The assumption that skilled workers are more likely to be successful entrepreneurs is supported by the evidence in Glaeser (2009). More skilled workers have $1 + H$ units of human capital, where $H > 0$, and have a value of $\alpha/(1+\alpha)$ equal to $\bar{\alpha}$. We assume that the high and low human capital workers are perfect substitutes in production and that the share of high human capital workers in city i is fixed at h_i (this is a closed-city model). In this case, the total number of employers is

$$E_i = \bar{E}_i - \Delta_i + [h_i\bar{\alpha} + (1 - h_i)\underline{\alpha}] N_i, \tag{31}$$

and the following result obtains.

Proposition 5 *If* $H\bar{E}_i/N_i + (1 + H)\underline{\alpha} > \bar{\alpha} > (1 + H)\underline{\alpha}$, *then there exists a value* $\bar{\Delta}_i \in (0, \bar{E}_i)$ *of the exogenous negative shock for which changes in human capital have no impact on the wages earned by each type of worker* ($\Delta_i = \bar{\Delta}_i \Leftrightarrow \partial w_i/\partial h_i = 0$). *If* Δ_i *is above that value wages rise with the share of skilled workers* ($\Delta_i > \bar{\Delta}_i \Leftrightarrow \partial w_i/\partial h_i > 0$), *and if* Δ_i *is below that value wages decline with the share of skilled workers* ($\Delta_i < \bar{\Delta}_i \Leftrightarrow \partial w_i/\partial h_i < 0$).

If $\bar{\alpha} \geq H\bar{E}_i/N_i + (1 + H)\underline{\alpha}$, *then wages for both classes of workers rise with the share of skilled workers* ($\partial w_i/\partial h_i \geq 0$ *for all* $\Delta_i \in [0, \bar{E}_i]$), *and if* $\bar{\alpha} \leq (1 + H)\underline{\alpha}$ *wages for both* *classes of workers fall with the share of the population that is skilled* ($\partial w_i / \partial h_i \leq 0$ *for all* $\Delta_i \in [0, \bar{E}_i]$*).*

Proposition 5 illustrates one way in which human-capital externalities might work. There are always two effects of having more skilled workers on earnings. More skilled workers can depress earnings because they are more productive and therefore lower the marginal product of labor when the number of employers is held fixed. But more skilled workers also increase the number of employers, and this causes wages to rise. If $\bar{\alpha}$ is higher than $H\bar{E}_i/N_i + (1 + H)\underline{\alpha}$, so skilled workers have a real comparative advantage at innovation, then wages will always rise with the share of skilled workers. This is one way in which human capital externalities might operate.

The proposition also illustrates the connection between adverse shocks and the value of having more skilled workers in the city. When there is more adverse economic shock that destroys the stock of old employers, then it is more likely that skilled workers will increase wages for everyone. When the shock is less severe, then skilled workers are less likely to improve everyone's welfare.

Proposition 5 examines the potential impact that skills can have on urban wages and success in the face of a downturn. The human capital needed to innovate might also result from experience in management, especially of smaller firms. We will not formally model this, but just note that the human capital needed to develop new firms may come from working in smaller, more entrepreneurial ventures. This would then be another reason why smaller firms are a source of urban resilience.

4 Why Do Educated Cities Grow?

We now turn to the primary statistical exercise of this paper: an examination of the link between education and metropolitan growth. Since we are focusing entirely on this later period, we switch from counties to metropolitan areas to be in line with past research. We also use data from entire United States. We follow Shapiro (2006) and Glaeser and Saiz (2004) and attempt to assess the reasons why skilled cities might grow more quickly. We differ from these earlier studies in two primary ways. First, we estimate all of our results for different regions. This enables us to estimate whether human capital has different effects in declining areas (e.g. the Midwest) and growing areas. Second, we use the methodology described in Section 3, which enables us to assess whether human capital is increasing population growth because of increasing productivity (or entrepreneurship), amenities or housing supply.

One set of regressions focus on metropolitan area level regressions, where our basic method is to regress:

$$\log \frac{Y_{2000}}{Y_{1970}} = B_Y \cdot \text{Schooling}_{1970} + \text{Other Controls.} \tag{32}$$

In this case, Y denotes one of three outcome variables: population, median income and self-reported housing values. We focus only on the long difference between 1970 and 2000. Our second approach is to use individual data and estimate:

$$\log Y_t = \text{MSA Dummies} + \text{Individual Controls} + B_Y \cdot \text{Schooling}_{1970} \cdot I_{2000}, \tag{33}$$

where Y in this case indicates either labor-market earnings or self-reported housing values. We pool together data for 1970 and 2000. In the case of the earnings regressions, individual controls include individual schooling, age and race. In the case of the housing value regressions, individual controls include structural characteristics such as the number of bedrooms and bathrooms. In both cases, we allow the coefficients on these characteristics to change by year and we include an indicator variable that takes on a value of one if the year is 2000.

Our primary focus is on the coefficient B_Y that multiplies the interaction between the share of the adult population with college degrees in 1970 and the year 2000. Essentially, this coefficient is assessing the extent to which housing values and incomes increased in more educated places. We prefer this specification to the raw income growth or housing value growth regressions because these regressions can control for differences in the returns to various individual characteristics.

One novelty of our work here is that we estimate the impact of education separately by regions. To do this, we interact B_Y with four region dummies, and thereby allow the impact of schooling on population, income and housing value growth to differ by region. These different regional parameter estimates will then imply different estimates of the underlying parameters found using the formulas of the last section.

Table 9 shows our results for metropolitan area level regressions. In all regressions, we include the initial values of the logarithm of population, median income and housing values. We also include three region dummies (the Midwest is the omitted category). The first regression shows the overall impact of education in this sample. As the share of the adult population with college degrees increased by 5 percent in 1970, predicted growth between 1970 and 2000 increases by about 8 percent.

The other coefficients in the regression are generally unsurprising. Growth was faster in the South and the West. Gibrat's law holds and population is unrelated to population growth. Places with higher housing values actually grew faster, perhaps because their expensiveness reflected a higher level of local amenities. Places with higher incomes grew more slowly, perhaps reflecting the movement away from high-wage, manufacturing metropolitan areas.

The second regression allows the impact of education in 1970 to differ by region. The strongest effect appears in the South, where a 5 percent increase in share of adults with college degrees in 1970 is associated with 19 percent faster population growth. The second largest coefficient appears in the Northeast. In that region, the coefficient is about the national average, even though it is not statistically significant. The coefficient is slightly smaller in the Midwest, where a 5 percentage point increase in the share of adults with college degrees in 1970 is associated with a 6.5 percentage point predicted increase in population between 1970 and 2000. In this case, however, the coefficient is statistically significant. In the West, the impact of education on population growth is negative and insignificant.

Our third regression looks at median growth in income. Income mean reverts, but increases in high housing value areas, perhaps suggesting that wealthier people are moving to higher-amenity areas. Incomes rose by less in the West; the other region dummies are statistically insignificant. There is a strong positive effect of initial education levels, which reflects in part the returns to skill and the tendency of skilled people to move to already skilled areas. As the share of the population with college degrees in 1970 increased by 5 percent, median incomes increase by 4 percent more since then.

The fourth regression estimates different initial education by region. Education has a positive effect on income growth in all four regions. The biggest impact is in the West, where income growth increases by .07 log points as the share of the population with college degrees in 1970 increases by 5 percentage points. The smallest impact of education on income growth is in the Midwest, where the coefficient is less than half of that found in the West.

The fifth and sixth regressions turn to appreciation in median housing values. Housing values rose by more in more populous metropolitan areas. Prices increased somewhat less in initially higher-income areas, perhaps reflecting the mean reversion of income levels. Prices, however, did not themselves mean revert. The West had much more price appreciation than the other three regions. As the share of the population with college degrees in 1970 increased by 5 percentage points, housing values increased by about 4 percent more.

The sixth regression allows the impact of college education on housing-value growth to differ by region. In this case, we find a big positive effect in the West, and far smaller effects in all other regions. In the West, prices rose by more than 10 percent more as the share of the population with college degrees in 1970 increased by 5 percentage points. In the other regions, the impact of education is statistically insignificant and less than one-fifth of its impact in the West. It is notable that the region where education had its weakest impact on population growth is the area where it had its largest impact on housing-value growth.

This difference shows the value of examining the impact of education by region.

The seventh and eight regressions examine the connection between education and average firm size. In the model, the number of firms per worker reflects the number of entrepreneurs in the area. If education is associated with a greater increase in population than in average firm size, then it is also associated with an increase in the number of firms, which we interpret as an increase in the amount of in entrepreneurship. In the seventh regression, we estimate a coefficient of .8 on college graduation rates across the whole sample. This coefficient is substantially lower than the population growth coefficient, so this suggests, at least according to the logic of the model, that the number of entrepreneurs is growing more quickly in more educated areas.

In the eighth regression, we allow the coefficient on education to differ by region. The strongest effect is in the East; the weakest in the West. In both the East and the West, the coefficient on education is higher in the average firm size regression than in the population growth regression. The very strong coefficient on average firm size in the East appears to be driven by two types of metropolitan areas. There are some less educated metropolitan areas where firm size is dramatically decreasing, presumably because large plants all closing. There are also some well educated metropolitan areas, including Boston, where firm size is increasing dramatically, perhaps because of the dominance of certain big-firm industries, such as health care. In the West, more educated areas seem also to be moving into larger, rather than smaller, firm industries.

Table 10 turns to wages and housing values using individual-level data. We look at annual earnings and restrict our sample to prime-age males (between 25 and 55), who work at least 30 hours a week and over 40 weeks per year. These restrictions are meant to limit issues associated with being out of the labor force. We control for individual human-capital characteristics, including years of experience and education, and allow for the impact of these variables to differ by area. As such, these coefficients can be understood as the impact of skills on area income growth correcting for the movement of skilled people across places and the rise in the returns to skill. All regressions also control for the initial levels of income, population and housing values, just like the metropolitan area level regressions. We also have MSA dummies in each regression, controlling for the permanent income differences between places.

The first regression shows a raw coefficient of .557, which implies that as the share of college graduates in a metropolitan area in 1970 increases by 5 percentage points, earnings rise by .028 log points more over the next thirty years. Comparing this coefficient with the coefficient on education (.8) in regression 3 in Table 10 suggests that almost a third of the metropolitan-area coefficient is explained by the rise in returns to skill at the individual level and increased sorting across metropolitan areas. The second regression adds in industry dummies, and the coefficient drops to .442.

The third regression compares the impact of education at the area level with education at the industry level in 1970. In this case, we allow the MSA dummies to differ by year, so these effects should be understood as across industries but within metropolitan area. The cross-industry effect of education on income growth is also positive, but it is much weaker than the effect at the metropolitan area level.

Regressions 4 and 5 look the impact of the initial education level in the MSA-industry. We calculate the share of workers in that metropolitan area in that industry in 1970 with college degrees. We then control for MSA-year dummies and industry fixed effects in regression 4. We find that more skilled sectors are seeing faster wage growth. Regression 5 shows that this effect does not withstand allowing the industry effects, nationwide, to vary by year.

Regression 6 essentially duplicates regression 1 of the table allowing the coefficient on education to differ by region. In this case, however, unlike the metropolitan area level tables, we find that there are few significant regional differences. The coefficient is slightly higher in the Northeast, but the effects are generally quite similar and close to the national effect.

In regressions 7 and 8 we estimate housing price appreciation using individual-level housing data and controlling for individual housing characteristics. Regression 7 shows the overall national coefficient of 3.3. Regression 8 estimates different effects by region, and again shows that housing price appreciation has gone up faster in the West.

Table 11 then shows our estimated coefficients, using the formulas in Section 3: $\beta_j^\theta = -B_{Inc} + (1-\mu)\, B_{Val}$, $\beta_j^{\bar{L}} = B_{Pop} + B_{Inc} - B_{Val}$, and $\tilde{\beta}_j^E = B_{Pop} + \sigma B_{Inc}$. We also use the firm size effect to separate the impact of education on "productivity," $\beta_j^a = (\sigma B_{Inc} + B_{Siz}) / (\sigma - 1)$, from the impact of education on "entrepreneurship," $\beta_j^E = B_{Pop} - B_{Siz}$. These enable us to combine these coefficients and assess whether education is acting on housing supply, productivity or amenities. To implement these equations we use a value of .7 for μ, which is compatible with housing representing 30 percent of consumption. For σ, we use a value of 4, which corresponds to an average mark-up of 33 percent. Jaimovich and Floetotto (2008) present some support for this calibration, which only impacts on the estimated connection between skills and productivity growth.

The first five columns show results for the country and each region using only the metropolitan area level coefficients. Columns 6-10 show results using the metropolitan area estimates for population growth and the area-level estimates for income and housing price growth. The estimates show standard errors estimated by bootstrap. However, we believe that these standard errors substantially overstate the actual precision of these estimates, since they take into account only the error involved in our estimated parameters, not the possibility that our assumed parameters, and indeed the model itself, are at best noisy approximations of reality.

The first column shows a positive connection between productivity growth and skills everywhere. The national coefficient is about 5, meaning that as the share of the population with college degrees increase by 5 percent, the growth in the number of entrepreneurs over the next 30 years increases by 25 percent. The coefficient is somewhat higher in the South and somewhat lower in the West, but these differences are not statistically significant. Using these national metropolitan-area coefficients, we find that the impact of education on the growth of productivity, or entrepreneurship, is reasonably homogeneous across regions.

The second column shows results for amenity growth. In every region the coefficient is negative, suggesting that amenities have been shrinking rather than growing in skilled areas. This comes naturally out of the model because real wages have, according to our formulation, been shrinking in skilled places. Again, with the metropolitan area level coefficients, the impact of skills on amenities is fairly similar across regions. However, if housing were a larger share of consumption or if housing prices were actually proxying for the growth of all prices, then the real wage effect would be zero and hence the implied connection between skills and amenity growth would be zero as well.

The third column looks at the growth of housing supply. Overall, skills have been associated with increases in housing supply, but there are very substantial regional differences. In the South, there is an extremely strong implied relationship between skills and housing supply growth. In the West, the implied relationship is negative. These differences reflect the very different relationship between skills and population growth in the South and in the West. We think that in a richer model with a better developed construction sector, these effects would appear as a movement along a supply curve rather than an actual shift in the supply of housing, and that the differences between West and South could be explained, at least in part, by very different housing supply elasticities (as found by Saiz, forthcoming).

Columns 4 and 5 decompose the overall productivity effect into an effect associated with rising values of firm-level productivity (a_j, column 4) and rising levels of entrepreneurship (E_j, column 5). Column 1 is equal to three times column 4 plus column 5 ($\tilde{\beta}_j^E = (\sigma - 1)\,\beta_j^a + \beta_j^E$). In the first row, we find that education is significantly associated with increases in firm-level productivity and insignificantly associated with increases in entrepreneurship. Overall, the firm-level productivity coefficient is responsibility for 84 percent of the connection between education and total productivity.

The next rows show that in both the East and the West, we find an insignificant negative connection between area education and the entrepreneurship measure. In these areas, education has a strong and quite similar positive correlation with firm-level productivity. In the Midwest and the South, there is more of a positive correlation between area education and entrepreneurship growth, and the connection is statistically significant in the South.

Columns 6-10 show results using individual-level regressions for housing and income. In Column 6, the skills coefficient on entrepreneurship growth is smaller, reflecting the fact that the connection between skills and income growth is lower in the individual-level regressions. We believe that these estimates are more defensible. As in column 1, the connection between skills and entrepreneurship seems strongest in the South and weakest in the West. In this case, the gulf in estimated coefficients is much larger and statistically significant. Understanding this regional gap seems like an important topic for future research.

Column 7 shows the connection between skills and amenity growth. Overall, the estimated coefficient is positive, but it is negative in three out of four regions. Only in the West are skills positively associated with implied amenity growth, meaning that only in the West are skills associated with declines in real wages. In the other regions, skills are associated with rising real wages, which implies a decline in amenities. As discussed above, we do not take this implication all that seriously, because it is quite sensitive to assumptions about the connection between housing prices and the overall price level. Moreover, if unobserved skill levels are rising in skilled metropolitan areas, then the rise in real wages, and hence the implied decline in amenity levels, would also be somewhat illusory. We are more confident about the difference between regions—the rise in the value of amenities in skilled areas in the West—than we are about the overall sign in the rest of the nation.

Column 8 shows the land growth effects, which are positive everywhere but in the West. Just as in column 6, the West is the one region where skills seem associated with a decline in housing availability. In this case, the effect seems to be quite strong, statistically and economically; and indeed, the West is so powerful that it makes the estimated national coefficient negative. Housing supply has grown very little in skilled areas in the West, perhaps because educated Westerners have been particularly effective in pushing for limits on new construction.

Columns 9 and 10 show results when we break overall productivity in firm-level productivity and the number of entrepreneurs. The basic patterns are quite similar to the MSA-level coefficients. Overall, the impact of education on both variables is positive, but the effect is only statistically significant for the firm-level productivity variable, which accounts for the lion's share of the overall productivity effect of education. In the East and West, the estimated coefficient of entrepreneurs on area level education is negative, but not statistically significant. In the South, the coefficient is strongly positive.

Overall, this exercise leads to four main conclusions. First, the impact of education on productivity seems to be quite clear everywhere. Second, the growth of skilled places has far more to do with rising productivity than with amenity growth outside of the West, and indeed, amenity levels may have been declining in skilled areas. This conclusion echoes the findings of Shapiro (2006) and Glaeser and Saiz (2004). Third, skills seem to depress housing supply growth in the American West, and that is a substantial difference with other regions. This negative connection could reflect the ability and taste of skilled people for organizing to oppose new construction. Fourth, the connection between education and overall productivity growth does not, outside of the South, primarily reflect a connection between education and an increase in the number of entrepreneurs.

5 Education and Unemployment in the Great Recession

The previous section focused on the role that education played in mediating cities' ability to respond to the great shocks of the mid-20th century, but there has also been a more recent crisis. According to the National Bureau of Economic Research, a recession began in December 2007. Unemployment then rose significantly in 2008 and 2009, rising above 10 percent in October 2009. But while the recession impacted on all of America, it did not hit every place equally. In February 2010, the unemployment rate was over 20 percent in Merced, California, and over 15 percent in Detroit, Michigan. At the same time, the unemployment rate in Minneapolis, Minnesota, was 7.7 percent and in Boulder, Colorado, only 6.5 percent.[9]

Just as education predicted the ability of older, colder cities to survive the mid-20th century shocks, skills also predict the ability of cities today to weather the storm. Figure 7 shows the -.44 correlation between the share of adults with a college degree in a metropolitan area and the unemployment rate in that area as of January 2010.

Although educational attainment is negatively correlated with unemployment at the individual level, the city-level correlation is too high to be entirely due to composition effects. We construct a predicted unemployment rate based on the breakdown of city population by education level:

$$\text{Predicted Unemployment} = \sum_{\text{Groups}} \text{Share}^{\text{MSA}}_{\text{Group}} \cdot U^{\text{USA}}_{\text{Group}}, \tag{34}$$

where $\text{Share}^{\text{MSA}}_{\text{Group}}$ is the share of the adult labor force in each group in each metropolitan area

[9]U.S. Department of Labor, Bureau of Labor Statistics. "Metropolitan Area Employment and Unemployment – April 2010," news release, June 2, 2010. http://www.bls.gov/news.release/pdf/metro.pdf.

in 2000 (the latest date available with reliable data), and $U_{\text{Group}}^{\text{USA}}$ is the national unemployment rate for the group, which was 5.1 percent for those with college degrees, 17.6 percent for high school dropouts and 10.25 percent for the remainder.

Figure 8 shows the .48 correlation between actual unemployment and our predicted unemployment measure. The key finding is that the slope of the regression line is 1.78: as predicted unemployment falls by 5 percent, actual unemployment declines by almost 8 percent. Education accounts for a greater decline in city unemployment than the national relationship between education and unemployment would imply. This provides another piece of evidence suggesting the existence of human capital spillovers.

Many interpretations of this fact are possible. It might be a coincidence that unemployment rates were unusually low in highly educated areas. People who live in educated areas could be more skilled than their years of schooling suggest. This in turn might reflect sorting, but also human capital spillovers that enhance unobserved skill levels (Glaeser 1999). The model in section 3 emphasized that skilled workers are both employers and employees. Hence the strong negative effect of education on unemployment may reflect the ability of more skilled entrepreneurs to find opportunity in a downturn. Of course, this explanation is now merely a hypothesis and further work will be needed to determine if it is correct.

6 Conclusion

The regional history of the eastern United States is best understood not through a set of immutable laws, but as a progression of different eras during which local attributes waxed and waned in importance. Few if any growth patterns hold over the entire 150 year period: many relationships, such as Gibrat's Law, hold during some periods but not others. During some periods growth is faster in more populous places, and during others population moves to more sparsely populated areas. Warmth positively predicts growth during the late 19th and 20th centuries, but not during the early parts of the two centuries.

To us, these findings support the view that regional and urban change is best understood not as the application of time-invariant growth processes, but rather as a set of responses by people and firms to large-scale technological change. These responses are quite amenable to formal modeling, but only to formal models that respect the changing nature of transportation and other technologies. The 19th century was primarily agricultural, and the spread west reflected the value of gaining access to highly productive agricultural land. The Great Lakes were a magnet because they lowered otherwise prohibitive transport costs. During the late 19th century, America became increasingly industrial and the population moved to places that began the era with more population. Cities, like Detroit and Chicago, that had formed as hubs for transporting the wealth of American agriculture became centers for producing manufactured goods such as cars.

Finally, during the post-war era, transportation costs fell still further and the population de-concentrated. The Great Lakes declined and people moved to the Sun Belt. The older areas that were best placed to reinvent themselves had a heavy concentration of skills and a disproportionate number of small firms, which may be a proxy for the level of entrepreneurial human capital. Industry no longer created a strong reason for concentration in populated counties, but it was increasingly valuable to be around skilled people. Our model formally addressed reinvention in skilled areas.

When we examine the channels through which skills affect growth, we find that growth in labor demand was significantly higher in more skilled areas, at least outside of the West. But in the West, skilled areas appear to have experienced faster amenity growth, perhaps because skilled people located in areas that were inherently more attractive. Skills were positively correlated with housing supply growth in the Midwest and South, but strongly negatively associated with housing supply growth in the West.

We also examine whether skills impact on labor demand primarily by increasing the number of entrepreneurs, as measured by the number of establishments in an area, or by increasing average firm level productivity. We find that education positively predicts growth in the number of establishments, but that this effect is relatively modest. The bulk of the connection between skills and labor demand appears to reflect a positive link between skills and average firm productivity.

America has experienced dramatic changes over the past 200 years, and population change doesn't appear to follow any form of strict rule. There has been a great deal of population persistence in the eastern U.S., but population change has followed different patterns at different times. Over the past thirty years, skills and small firms have been strongly correlated with growth, but that may not always be the case.

A Appendix

A.1. Setup of the Model

The utility function (3) implies the demand function for each manufactured variety

$$q(\nu) = \mu Y P^{\sigma-1} p(\nu)^{-\sigma}, \tag{A1}$$

where Y is nominal aggregate income in the whole economy and

$$P \equiv \left[\int_0^G p(\nu)^{1-\sigma} \, d\nu \right]^{1/(1-\sigma)} \tag{A2}$$

is the manufacturing price index, which we can set equal to one by a choice of numeraire. If the wage in city i is w_i, the price charged by the monopolistically competitive producer of each good ν manufactured in city i equals

$$p(\nu) = \frac{\sigma}{\sigma-1} \frac{w_i}{a_i}, \tag{A3}$$

and labor demand from each manufacturer equals

$$n(\nu) = \left(\frac{\sigma-1}{\sigma} \right)^{\sigma} \mu Y a_i^{\sigma-1} w_i^{-\sigma}. \tag{A4}$$

With E_i producer labor demand in city i equals

$$N_i = \left(\frac{\sigma-1}{\sigma} \right)^{\sigma} \mu Y E_i a_i^{\sigma-1} w_i^{-\sigma}, \tag{A5}$$

which yields the equilibrium wage (5) for a given labor supply N_i. The profits earned by each entrepreneur are then

$$\pi_i = \frac{1}{\sigma} \left[\mu Y \left(a_i \frac{N_i}{E_i} \right)^{\sigma-1} \right]^{\frac{1}{\sigma}}. \tag{A6}$$

The spatial equilibrium condition (7) can be rewritten as

$$\frac{1}{N_i} \left[\theta_i^{\sigma} a_i^{\mu(\sigma-1)} E_i^{\mu} \bar{L}_i^{(1-\mu)\sigma} \right]^{\frac{1}{\mu+\sigma-\mu\sigma}} = \frac{1}{N_j} \left[\theta_j^{\sigma} a_j^{\mu(\sigma-1)} E_j^{\mu} \bar{L}_j^{(1-\mu)\sigma} \right]^{\frac{1}{\mu+\sigma-\mu\sigma}} \text{ for all } i, j, \tag{A7}$$

which implies equation (8) for

$$\kappa_N \equiv \log N - \log \int \left[\theta_j^{\sigma} a_j^{\mu(\sigma-1)} E_j^{\mu} \bar{L}_j^{(1-\mu)\sigma} \right]^{\frac{1}{\mu+\sigma-\mu\sigma}} dj. \tag{A8}$$

Aggregate income is

$$Y = \frac{1}{\psi\mu}\left\{\int\left[E_j\left(a_jN_j\right)^{\sigma-1}\right]^{\frac{1}{\sigma}}dj\right\}^{\frac{\sigma}{\sigma-1}}$$

$$= \frac{N}{\psi\mu}\frac{\left\{\int\left[\left(\theta_ja_j\right)^{\sigma-1}E_j\bar{L}_j^{(1-\mu)(\sigma-1)}\right]^{\frac{1}{\mu+\sigma-\mu\sigma}}dj\right\}^{\frac{\sigma}{\sigma-1}}}{\int\left[\theta_j^{\sigma}a_j^{\mu(\sigma-1)}E_j^{\mu}\bar{L}_j^{(1-\mu)\sigma}\right]^{\frac{1}{\mu+\sigma-\mu\sigma}}dj}, \quad \text{(A9)}$$

so equation (9) holds for

$$\kappa_w \equiv \log\frac{\sigma-1}{\sigma} - \log\psi + \frac{1}{\sigma-1}\log\int\left[\left(\theta_ja_j\right)^{\sigma-1}E_j\bar{L}_j^{(1-\mu)(\sigma-1)}\right]^{\frac{1}{\mu+\sigma-\mu\sigma}}dj, \quad \text{(A10)}$$

and equation (10) for

$$\kappa_r \equiv \log\frac{1-\mu}{\mu} - \log\frac{\sigma-1}{\sigma} + \kappa_N + \kappa_w. \quad \text{(A11)}$$

A.2. Proof of Proposition 1

Equations (8), (9), and (10) yield

$$\frac{\partial Var\left(\log N_{i,t}\right)}{\partial Var\left(\Gamma_i\right)} = \left[\frac{\mu\left(\sigma-1\right)}{\left(\mu+\sigma-\mu\sigma\right)T_t}\right]^2, \quad \text{(A12)}$$

so

$$\frac{\partial^2 Var\left(\log N_{i,t}\right)}{\partial T_t\partial Var\left(\Gamma_i\right)} = -\frac{2}{T_t^3}\left[\frac{\mu\left(\sigma-1\right)}{\mu+\sigma-\mu\sigma}\right]^2 < 0; \quad \text{(A13)}$$

$$\frac{\partial Var\left(\log w_{i,t}\right)}{\partial Var\left(\Gamma_i\right)} = \left[\frac{\left(1-\mu\right)\left(\sigma-1\right)}{\left(\mu+\sigma-\mu\sigma\right)T_t}\right]^2, \quad \text{(A14)}$$

so

$$\frac{\partial^2 Var\left(\log w_{i,t}\right)}{\partial T_t\partial Var\left(\Gamma_i\right)} = -\frac{2}{T_t^3}\left[\frac{\left(1-\mu\right)\left(\sigma-1\right)}{\mu+\sigma-\mu\sigma}\right]^2 < 0; \quad \text{(A15)}$$

and

$$\frac{\partial Var\left(\log r_{i,t}\right)}{\partial Var\left(\Gamma_i\right)} = \left[\frac{\sigma-1}{\left(\mu+\sigma-\mu\sigma\right)T_t}\right]^2, \quad \text{(A16)}$$

so

$$\frac{\partial^2 Var\left(\log r_{i,t}\right)}{\partial T_t\partial Var\left(\Gamma_i\right)} = -\frac{2}{T_t^3}\left(\frac{\sigma-1}{\mu+\sigma-\mu\sigma}\right)^2 < 0; \quad \text{(A17)}$$

while for all $\Xi_{i,t} \in \{N_{i,t}, w_{i,t}, r_{i,t}\}$,

$$\frac{\partial^2 Var\left(\log \Xi_{i,t}\right)}{\partial T_t\partial Var\left(\theta_{i,t}\right)} = \frac{\partial^2 Var\left(\log \Xi_{i,t}\right)}{\partial T_t\partial Var\left(A_{i,t}\right)} = \frac{\partial^2 Var\left(\log \Xi_{i,t}\right)}{\partial T_t\partial Var\left(E_{i,t}\right)} = \frac{\partial^2 Var\left(\log \Xi_{i,t}\right)}{\partial T_t\partial Var\left(\bar{L}_{i,t}\right)} = 0. \quad \text{(A18)}$$

A.3. Proof of Proposition 2

Equation (27) and (28) imply that the market-clearing wage is

$$w_i = \frac{\sigma-1}{\sigma}\left\{\mu Y \frac{a_i^{\sigma-1}}{1-t\left(1-\bar{\eta}_i^{\alpha}\right)}\left[\frac{\bar{E}_i-\Delta_i}{M_i}+\frac{\alpha}{1+\alpha}\left(1-\bar{\eta}_i^{1+\alpha}\right)\right]\right\}^{\frac{1}{\sigma}}, \tag{A19}$$

and the profits of each successful entrepreneur are

$$\pi_i = \frac{1}{\sigma}\left(\mu Y\right)^{\frac{1}{\sigma}}\left\{\frac{1}{a_i\left[1-t\left(1-\bar{\eta}_i^{\alpha}\right)\right]}\left[\frac{\bar{E}_i-\Delta_i}{M_i}+\frac{\alpha}{1+\alpha}\left(1-\bar{\eta}_i^{1+\alpha}\right)\right]\right\}^{\frac{1-\sigma}{\sigma}}. \tag{A20}$$

In particular if $t=1$, so people are either would-be entrepreneurs or employees, then

$$\bar{\eta}_i = \begin{cases} 1 & \text{if } M_i \le (\sigma-1)\left(\bar{E}_i-\Delta_i\right) \\ \left\{\frac{\sigma-1}{1+\alpha\sigma}\left[(1+\alpha)\frac{\bar{E}_i-\Delta_i}{M_i}+\alpha\right]\right\}^{\frac{1}{1+\alpha}} & \text{if } M_i \ge (\sigma-1)\left(\bar{E}_i-\Delta_i\right), \end{cases} \tag{A21}$$

the total number of employers equals

$$E_i = \begin{cases} \bar{E}_i-\Delta_i & \text{if } M_i \le (\sigma-1)\left(\bar{E}_i-\Delta_i\right) \\ \frac{(1+\alpha)\left(\bar{E}_i-\Delta_i\right)+\alpha M_i}{1+\alpha\sigma} & \text{if } M_i \ge (\sigma-1)\left(\bar{E}_i-\Delta_i\right), \end{cases} \tag{A22}$$

and wages are

$$w_i = \begin{cases} \frac{\sigma-1}{\sigma}\left[\mu Y\left(\frac{\tau_i}{\psi}\right)^{\sigma-1}\frac{\bar{E}_i-\Delta_i}{M_i}\right]^{\frac{1}{\sigma}} & \text{if } M_i \le (\sigma-1)\left(\bar{E}_i-\Delta_i\right) \\ \frac{\sigma-1}{\sigma}\left\{\mu Y\left(\frac{\tau_i}{\psi}\right)^{\sigma-1}\left[\frac{\alpha+(1+\alpha)\left(\bar{E}_i-\Delta_i\right)/M_i}{(1+\alpha\sigma)(\sigma-1)^{\alpha}}\right]^{\frac{1}{1+\alpha}}\right\}^{\frac{1}{\sigma}} & \text{if } M_i \ge (\sigma-1)\left(\bar{E}_i-\Delta_i\right). \end{cases} \tag{A23}$$

The response of wages to a negative shock is

$$\frac{\partial \log w_i}{\partial \Delta_i} = \left\{\begin{array}{ll} -\left[\sigma\left(\bar{E}_i-\Delta_i\right)\right]^{-1} & \text{if } \Delta_i < \bar{E}_i-\frac{M_i}{\sigma-1} \\ -\left\{\sigma\left[\alpha M_i+(1+\alpha)\left(\bar{E}_i-\Delta_i\right)\right]\right\}^{-1} & \text{if } \Delta_i > \bar{E}_i-\frac{M_i}{\sigma-1} \end{array}\right\} < 0, \tag{A24}$$

such that

$$\frac{\partial^2 \log w_i}{\partial\alpha\partial\Delta_i} = \left\{\begin{array}{ll} 0 & \text{if } \Delta_i < \bar{E}_i-\frac{M_i}{\sigma-1} \\ \frac{M_i+\bar{E}_i-\Delta_i}{\sigma\left[\alpha M_i+(1+\alpha)\left(\bar{E}_i-\Delta_i\right)\right]^2} & \text{if } \Delta_i > \bar{E}_i-\frac{M_i}{\sigma-1} \end{array}\right\} \ge 0, \tag{A25}$$

with a convex kink at $\Delta_i = \bar{E}_i - M_i/(\sigma-1)$.

The number of entrepreneurs reacts according to

$$\frac{\partial \log E_i}{\partial \Delta_i} = \left\{\begin{array}{ll} -\left(\bar{E}_i-\Delta_i\right)^{-1} & \text{if } \Delta_i < \bar{E}_i-\frac{M_i}{\sigma-1} \\ -\left(\bar{E}_i-\Delta_i+\frac{\alpha}{1+\alpha}M_i\right)^{-1} & \text{if } \Delta_i > \bar{E}_i-\frac{M_i}{\sigma-1} \end{array}\right\} < 0, \tag{A26}$$

such that

$$\frac{\partial^2 \log E_i}{\partial\alpha\partial\Delta_i} = \left\{ \begin{array}{ll} 0 & \text{if} \quad \Delta_i < \bar{E}_i - \frac{M_i}{\sigma-1} \\ M_i\left[(1+\alpha)\left(\bar{E}_i - \Delta_i\right) + \alpha M_i\right]^{-2} & \text{if} \quad \Delta_i > \bar{E}_i - \frac{M_i}{\sigma-1} \end{array} \right\} \geq 0, \tag{A27}$$

with a convex kink at $\Delta_i = \bar{E}_i - M_i/(\sigma - 1)$.

A.4. Proof of Proposition 3

Average firm size is

$$n_i = \left\{ \begin{array}{ll} \frac{\bar{M}}{\bar{E}_i - \Delta_i} & \text{if} \quad \Delta_i \leq \bar{E}_i - \frac{\bar{M}}{\sigma-1} \\ \left[\frac{(1+\alpha\sigma)(\sigma-1)^\alpha}{\alpha+(1+\alpha)\left(\bar{E}_i-\Delta_i\right)/\bar{M}}\right]^{\frac{1}{1+\alpha}} & \text{if} \quad \Delta_i \geq \bar{E}_i - \frac{\bar{M}}{\sigma-1}, \end{array} \right. \tag{A28}$$

such that

$$\frac{\partial n_i}{\partial\left(\bar{E}_i - \Delta_i\right)} = \left\{ \begin{array}{ll} -\bar{M}\left(\bar{E}_i - \Delta_i\right)^{-2} & \text{if} \quad \Delta_i \leq \bar{E}_i - \frac{\bar{M}}{\sigma-1} \\ -\frac{\left[(1+\alpha\sigma)(\sigma-1)^\alpha\right]^{\frac{1}{1+\alpha}}}{\bar{M}}\left[\alpha + (1+\alpha)\frac{\bar{E}_i-\Delta_i}{\bar{M}}\right]^{-\frac{2+\alpha}{1+\alpha}} & \text{if} \quad \Delta_i \geq \bar{E}_i - \frac{\bar{M}}{\sigma-1} \end{array} \right. < 0. \tag{A29}$$

The response of wages to a negative shock has

$$\frac{\partial^2 \log w_i}{\partial\Delta_i\partial\left(\bar{E}_i - \Delta_i\right)} = \left\{ \begin{array}{ll} \frac{1}{\sigma}\left(\bar{E}_i - \Delta_i\right)^{-2} & \text{if} \quad \Delta_i < \bar{E}_i - \frac{\bar{M}}{\sigma-1} \\ \frac{1+\alpha}{\sigma}\left[\alpha\bar{M} + (1+\alpha)\left(\bar{E}_i - \Delta_i\right)\right]^{-2} & \text{if} \quad \Delta_i > \bar{E}_i - \frac{\bar{M}}{\sigma-1} \end{array} \right\} > 0, \tag{A30}$$

which implies

$$\frac{\partial^2 \log w_i}{\partial\Delta_i\partial n_i} = \left\{ \begin{array}{ll} -\frac{1}{\sigma\bar{M}} & \text{if} \quad \Delta_i < \bar{E}_i - \frac{\bar{M}}{\sigma-1} \\ -\frac{1+\alpha}{(1+\alpha\sigma)\sigma(\sigma-1)^\alpha}\frac{n_i^\alpha}{\bar{M}} & \text{if} \quad \Delta_i > \bar{E}_i - \frac{\bar{M}}{\sigma-1} \end{array} \right\} < 0. \tag{A31}$$

The number of entrepreneurs reacts according to

$$\frac{\partial^2 \log E_i}{\partial\Delta_i\partial\left(\bar{E}_i - \Delta_i\right)} = \left\{ \begin{array}{ll} \left(\bar{E}_i - \Delta_i\right)^{-2} & \text{if} \quad \Delta_i < \bar{E}_i - \frac{\bar{M}}{\sigma-1} \\ \left(\bar{E}_i - \Delta_i + \frac{\alpha}{1+\alpha}\bar{M}\right)^{-2} & \text{if} \quad \Delta_i > \bar{E}_i - \frac{\bar{M}}{\sigma-1} \end{array} \right\} > 0, \tag{A32}$$

which implies

$$\frac{\partial^2 \log E_i}{\partial\Delta_i\partial n_i} = \left\{ \begin{array}{ll} -\frac{1}{\bar{M}} & \text{if} \quad \Delta_i < \bar{E}_i - \frac{\bar{M}}{\sigma-1} \\ -\frac{(1+\alpha)^2}{(1+\alpha\sigma)(\sigma-1)^\alpha}\frac{n_i^\alpha}{\bar{M}} & \text{if} \quad \Delta_i > \bar{E}_i - \frac{\bar{M}}{\sigma-1} \end{array} \right\} < 0. \tag{A33}$$

A.5. Proof of Proposition 4

For $t = 0$, the total number of employers equals

$$E_i = \bar{E}_i - \Delta_i + \frac{\alpha}{1+\alpha} N_i, \tag{A34}$$

wages are

$$w_i = \frac{\sigma - 1}{\sigma} \left[\mu Y a_i^{\sigma-1} \left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \right]^{\frac{1}{\sigma}}, \tag{A35}$$

and the profits of a successful entrepreneur are

$$\pi_i = \frac{1}{\sigma} \left[\mu Y a_i^{\sigma-1} \left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right)^{1-\sigma} \right]^{\frac{1}{\sigma}}, \tag{A36}$$

so expected earnings equal

$$\begin{aligned} y_i &\equiv w_i + \frac{\alpha}{1+\alpha} \pi_i \\ &= \left(\frac{\sigma - 1}{\sigma} \frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \left[\mu Y a_i^{\sigma-1} \left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right)^{1-\sigma} \right]^{\frac{1}{\sigma}}, \end{aligned} \tag{A37}$$

with

$$\begin{aligned} &\frac{\partial \log y_i}{\partial \left[\left(\bar{E}_i - \Delta_i \right) / N_i \right]} \\ &\quad = \frac{\sigma - 1}{\sigma^2} \frac{\bar{E}_i - \Delta_i}{N_i} \left[\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \left(\frac{\sigma - 1}{\sigma} \frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \right]^{-1} > 0, \end{aligned} \tag{A38}$$

so that $\partial \log y_i / \partial N_i < 0$, $\partial \log y_i / \partial \bar{E}_i > 0$ and$\partial \log y_i / \partial \Delta_i < 0$, while

$$\begin{aligned} \frac{\partial \log y_i}{\partial \alpha} &= \frac{1}{(1+\alpha)^2 \sigma} \left(\frac{2\sigma - 1}{\sigma} \frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \\ &\quad \cdot \left[\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \left(\frac{\sigma - 1}{\sigma} \frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \right]^{-1} > 0. \end{aligned} \tag{A39}$$

The price of land is

$$r_i = (1 - \mu) \left[Y \left(\frac{a_i}{\mu} \right)^{\sigma - 1} \left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \right]^{\frac{1}{\sigma}} \frac{N_i}{\bar{L}_i}, \tag{A40}$$

with

$$\frac{\partial \log r_i}{\partial N_i} = \frac{1}{N_i} \left(\frac{\sigma - 1}{\sigma} \frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right) \left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha} \right)^{-1} > 0, \tag{A41}$$

and

$$\frac{\partial \log r_i}{\partial \left(\bar{E}_i - \Delta_i\right)} = \frac{1}{\sigma N_i}\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)^{-1} > 0, \tag{A42}$$

so that $\partial \log r_i / \partial \bar{E}_i > 0$ and $\partial \log r_i / \partial \Delta_i < 0$, while

$$\frac{\partial \log r_i}{\partial \alpha} = \frac{1}{\left(1+\alpha\right)^2 \sigma}\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)^{-1} > 0. \tag{A43}$$

The spatial equilibrium requirement $\theta_i y_i r_i^{\mu-1} = \bar{U}$ can be written

$$\theta_i a_i^{\mu\frac{\sigma-1}{\sigma}}\left(\frac{\bar{L}_i}{N_i}\right)^{1-\mu}\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)^{\frac{\mu-\sigma}{\sigma}} = \left(1-\mu\right)^{1-\mu}\mu^{\frac{\mu\sigma-\mu-\sigma}{\sigma}}Y^{-\frac{\mu}{\sigma}}\bar{U}. \tag{A44}$$

With a continuum of cities, changes in a single atomistic city i do not affect the aggregate variables on the right-hand side, so the effects of changes in α and $\bar{E}_i - \Delta_i$ on the equilibrium workforce N_i can be taken from the constancy of

$$\Omega \equiv \left(1-\mu\right)\log N_i + \frac{\sigma-\mu}{\sigma}\log\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right) - \log\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right), \tag{A45}$$

such that

$$\frac{\partial \Omega}{\partial N_i} = \frac{1}{N_i}\left[\frac{\sigma-1}{\sigma^2}\left(\frac{\bar{E}_i - \Delta_i}{N_i}\right)^2 + \left(1-\mu\right)\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)^2\right] \cdot \left[\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)\right]^{-1} > 0, \tag{A46}$$

$$\frac{\partial \Omega}{\partial \alpha} = -\frac{1}{\left(1+\alpha\right)^2}\left[\frac{\sigma-\mu+\mu\sigma}{\sigma^2}\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha\mu}{\left(1+\alpha\right)\sigma}\right] \cdot \left[\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)\right]^{-1} < 0, \tag{A47}$$

and

$$\frac{\partial \Omega}{\partial \left(\bar{E}_i - \Delta_i\right)} = \frac{1}{\sigma N_i}\left[\frac{\alpha\left(1-\mu\right)}{1+\alpha} - \frac{\mu\left(\sigma-1\right)}{\sigma}\frac{\bar{E}_i - \Delta_i}{N_i}\right] \cdot \left[\left(\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i - \Delta_i}{N_i} + \frac{\alpha}{1+\alpha}\right)\right]^{-1}, \tag{A48}$$

which switches from positive to negative as μ ranges in $(0, 1)$.

The exogenous endowment of entrepreneurs has an ambiguous impact on the labor supply:

$$\frac{\partial N_i}{\partial\left(\bar{E}_i-\Delta_i\right)}=\frac{1}{\sigma}\left[\frac{\mu(\sigma-1)}{\sigma}\frac{\bar{E}_i-\Delta_i}{N_i}-\frac{\alpha(1-\mu)}{1+\alpha}\right]$$
$$\cdot\left[\frac{\sigma-1}{\sigma^2}\left(\frac{\bar{E}_i-\Delta_i}{N_i}\right)^2+(1-\mu)\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i-\Delta_i}{N_i}+\frac{\alpha}{1+\alpha}\right)^2\right]^{-1}, \quad \text{(A49)}$$

such that $\partial^2 N_i/\left[\partial\mu\partial\left(\bar{E}_i-\Delta_i\right)\right]>0$ for all $\mu\in(0,1)$ and

$$-\frac{\alpha}{1+\alpha}\frac{1}{\sigma}\left[\frac{\sigma-1}{\sigma^2}\left(\frac{\bar{E}_i-\Delta_i}{N_i}\right)^2+\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i-\Delta_i}{N_i}+\frac{\alpha}{1+\alpha}\right)^2\right]^{-1}$$
$$\leq\frac{\partial N_i}{\partial\left(\bar{E}_i-\Delta_i\right)}\leq\frac{N_i}{\bar{E}_i-\Delta_i}. \quad \text{(A50)}$$

Taking into account the endogenous response of the workforce, expected earnings are decreasing in Δ_i:

$$\frac{d\log y_i}{d\left(\bar{E}_i-\Delta_i\right)}=\frac{\partial\log y_i}{\partial\left(\bar{E}_i-\Delta_i\right)}+\frac{\partial\log y_i}{\partial N_i}\frac{\partial N_i}{\partial\left(\bar{E}_i-\Delta_i\right)}$$
$$\geq\frac{\partial\log y_i}{\partial\left(\bar{E}_i-\Delta_i\right)}+\frac{\partial\log y_i}{\partial N_i}\frac{N_i}{\bar{E}_i-\Delta_i}=0, \quad \text{(A51)}$$

and so are the total number of employers

$$\frac{dE_i}{d\left(\bar{E}_i-\Delta_i\right)}=1+\frac{\alpha}{1+\alpha}\frac{\partial N_i}{\partial\left(\bar{E}_i-\Delta_i\right)}$$
$$\geq\frac{\sigma-1}{\sigma}\left(\frac{\bar{E}_i-\Delta_i}{N_i}+\frac{\alpha}{1+\alpha}\right)^2\left[\frac{\sigma-1}{\sigma^2}\left(\frac{\bar{E}_i-\Delta_i}{N_i}\right)^2+\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i-\Delta_i}{N_i}+\frac{\alpha}{1+\alpha}\right)^2\right]^{-1}$$
$$>0 \quad \text{(A52)}$$

and the price of land

$$\frac{d\log r_i}{d\left(\bar{E}_i-\Delta_i\right)}=\frac{\partial\log r_i}{\partial\left(\bar{E}_i-\Delta_i\right)}+\frac{\partial\log r_i}{\partial N_i}\frac{\partial N_i}{\partial\left(\bar{E}_i-\Delta_i\right)}$$
$$\geq\frac{\sigma-1}{\sigma^2}\frac{\bar{E}_i-\Delta_i}{N_i^2}\left[\frac{\sigma-1}{\sigma^2}\left(\frac{\bar{E}_i-\Delta_i}{N_i}\right)^2+\left(\frac{\sigma-1}{\sigma}\frac{\bar{E}_i-\Delta_i}{N_i}+\frac{\alpha}{1+\alpha}\right)^2\right]^{-1}>0. \quad \text{(A53)}$$

City population equals $\Lambda_i\equiv\bar{E}_i-\Delta_i+N_i$, which is increasing in $\bar{E}_i-\Delta_i$ if, but not only if,

$$\sigma\geq\frac{1+\alpha}{\alpha}\geq 2\Leftrightarrow\alpha\geq\frac{1}{\sigma-1}. \quad \text{(A54)}$$

The elasticity of endogenous entrepreneurship α increases the labour supply

$$\frac{\partial N_i}{\partial \alpha} = -\frac{\partial \Omega / \partial \alpha}{\partial \Omega / \partial N_i} > 0, \tag{A55}$$

and therefore population, as well as the total number of employers

$$\frac{dE_i}{d\alpha} = \frac{N_i}{(1+\alpha)^2} + \frac{\alpha}{1+\alpha}\frac{\partial N_i}{\partial \alpha} > 0, \tag{A56}$$

land prices

$$\frac{d \log r_i}{d\alpha} = \frac{\partial \log r_i}{\partial \alpha} + \frac{\partial \log r_i}{\partial N_i}\frac{\partial N_i}{\partial \alpha} > 0, \tag{A57}$$

and expected earnings

$$\frac{d \log y_i}{d\alpha} = \frac{\partial \log y_i}{\partial \alpha} + \frac{\partial \log y_i}{\partial N_i}\frac{\partial N_i}{\partial \alpha} > 0, \tag{A58}$$

which can be verified with tedious but straightforward algebra.

A.6. Proof of Proposition 5

The wage per effective unit of human capital equals

$$w_i = \frac{\sigma - 1}{\sigma}\left\{\frac{\mu Y a_i^{\sigma-1}}{1+h_i H}\left[\frac{\bar{E}_i - \Delta_i}{N_i} + h_i \bar{\alpha} + (1-h_i)\underline{\alpha}\right]\right\}^{\frac{1}{\sigma}}, \tag{A59}$$

such that

$$\frac{\partial \log w_i}{\partial h_i} = \frac{1}{\sigma}\left[\bar{\alpha} - (1+H)\underline{\alpha} - H\frac{\bar{E}_i - \Delta_i}{N_i}\right]\left\{(1+h_i H)\left[\frac{\bar{E}_i - \Delta_i}{N_i} + h_i \bar{\alpha} + (1-h_i)\underline{\alpha}\right]\right\}^{-1}. \tag{A60}$$

Thus

$$\bar{\alpha} \geq (1+H)\underline{\alpha} + H\frac{\bar{E}_i}{N_i} \Rightarrow \frac{\partial \log w_i}{\partial h_i} \geq 0 \text{ for all } \Delta_i \in \left[0, \bar{E}_i\right], \tag{A61}$$

and

$$\bar{\alpha} \leq (1+H)\underline{\alpha} \Rightarrow \frac{\partial \log w_i}{\partial h_i} \leq 0 \text{ for all } \Delta_i \in \left[0, \bar{E}_i\right], \tag{A62}$$

while if $H\bar{E}_i/N_i + (1+H)\underline{\alpha} > \bar{\alpha} > (1+H)\underline{\alpha}$, then

$$\frac{\partial \log w_i}{\partial h_i} = 0 \Leftrightarrow \Delta_i = \bar{E}_i - \frac{\bar{\alpha} - (1+H)\underline{\alpha}}{H}N_i \equiv \bar{\Delta}_i, \tag{A63}$$

and wages are increasing in h_i for $\Delta_i > \bar{\Delta}_i$ and decreasing in h_i for $\Delta_i < \bar{\Delta}_i$.

References

Audretsch, David B. 1995. *Innovation and Industry Evolution.* Cambridge, Mass.: MIT Press.

———. 2003. Entrepreneurship: A Survey of the Literature. Enterprise Paper No. 14. Enterprise Directorate-General, European Commission, Brussels.

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

Berry, Christopher R., and Edward L. Glaeser. 2005. The Divergence of Human Capital Levels across Cities. *Papers in Regional Science*, 84(3): 407–444.

Black, Duncan, and J. Vernon Henderson. 1999. A Theory of Urban Growth. *Journal of Political Economy*, 107(2): 252–284.

Borjas, George J. 2003. The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *Quarterly Journal of Economics*, 118(4): 1335–1374.

Córdoba, Juan-Carlos. 2008. On the Distribution of City Sizes. *Journal of Urban Economics*, 63(1): 177–197.

Desmet, Klaus, and Esteban Rossi-Hansberg. 2009a. Spatial Growth and Industry Age. *Journal of Economic Theory*, 144(6): 2477–2502.

———. 2009b. Spatial Development. NBER Working Paper No. 15349. National Bureau of Economic Research, Cambridge, Mass.

———. 2010. On Spatial Dynamics. *Journal of Regional Science*, 50(1): 43–63.

Duranton, Gilles. 2006. Some Foundations for Zipf's Law: Product Proliferation and Local Spillovers. *Regional Science and Urban Economics*, 36(4): 543–563.

———. 2007. Urban Evolutions: The Fast, the Slow, and the Still. *American Economic Review*, 97(1): 197–221.

Eaton, Jonathan, and Zvi Eckstein. 1997. Cities and Growth: Theory and Evidence from France and Japan. *Regional Science and Urban Economics*, 27: 443–474.

Eeckhout, Jan. 2004. Gibrat's Law for (All) Cities. *American Economic Review*, 94(5): 1429–1451.

Farley, Reynolds. 2008. Census Taking and Census Undercount: Prickly Statistical, Political and Constitutional Issues. Paper presented at the National Poverty Center 2008 Summer Workshop: Analyzing Poverty and Socioeconomic Trends Using the American Community Survey, June 23–27, in Ann Arbor, Michigan.

Gabaix, Xavier. 1999. Zipf's Law for Cities: An Explanation. *Quarterly Journal of Economics*, 114(3): 739–767.

Gabaix, Xavier, and Yannis M. Ioannides. 2004. The Evolution of City Size Distributions. In *Handbook of Regional and Urban Economics, Vol. 4: Cities and Geography*, ed. J. V. Henderson and J.-F. Thisse, 2341–2378. Amsterdam: Elsevier.

Glaeser, Edward L. 1999. Learning in Cities. *Journal of Urban Economics*, 46: 254–277.

———. 2009. Entrepreneurship and the City. In *Entrepreneurship and Openness: Theory and Evidence*, ed. D. B. Audretsch, R. Litan, and R. Strom, 131–180. Cheltenham: Edward Elgar.

Glaeser, Edward L., and Joseph Gyourko. 2005. Urban Decline and Durable Housing. *Journal of Political Economy*, 113(2): 345–75.

Glaeser, Edward L., Joseph Gyourko, and Raven E. Saks. 2006. Urban growth and housing supply. *Journal of Economic Geography*, 6(1): 71–89.

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

Glaeser, Edward L., William R. Kerr, and Giacomo A. M. Ponzetto. 2010. Clusters of Entrepreneurship. *Journal of Urban Economics*, 67(1): 150–168.

Glaeser, Edward L., and Janet E. Kohlhase, 2004. Cities, Regions and the Decline of Transport Costs. *Papers in Regional Science*, 83(1): 197–228.

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

Glaeser, Edward L., and Matthew G. Resseger. 2010. The Complementarity between Cities and Skills. *Journal of Regional Science*, 50(1): 221–244.

Glaeser, Edward L., and Albert Saiz. 2004. The Rise of the Skilled City. *Brookings-Wharton Papers on Urban Affairs*, 5: 47–94.

Glaeser, Edward L., José A. Scheinkman, and Andrei Shleifer. 1995. Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economics*, 36(1): 117–143.

Glaeser, Edward L., and Jesse M. Shapiro. 2003. Urban Growth in the 1990s: Is City Living Back? *Journal of Regional Science*, 43(1): 139–165.

Glaeser, Edward L., and Kristina Tobio. 2008. The Rise of the Sunbelt. *Southern Economic Journal*, 74(3): 610–643.

Haines, Michael R., and Inter-university Consortium for Political and Social Research. 2010. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. Inter-university Consortium for Political and Social Research, Ann Arbor, Mich.

Holmes, Thomas J. 1998. The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders. *Journal of Political Economy*, 106(4): 667–705.

Jaimovich, Nir, and Max Floetotto. 2008. Firm Dynamics, Markup Variations and the Business Cycle. *Journal of Monetary Economics*, 55(7): 1238–1252.

Kim, Sukkoo. 2006. Division of Labor and the Rise of Cities: Evidence from US Industrialization, 1850–1880. *Journal of Economic Geography*, 6: 469–491.

Kim, Sukkoo, and Robert A. Margo. 2004. Historical Perspectives on U.S. Economic Geography. In *Handbook of Regional and Urban Economics, Vol. 4: Cities and Geography*, ed. J. V. Henderson and J.-F. Thisse, 2981–3019. Amsterdam: Elsevier.

Miracky, William F. 1995. Economic Growth in Cities: The Role of Localization Externalities. PhD diss., M.I.T.

Moretti, Enrico. 2004. Estimating the Social Return to Higher Education: Evidence From Cross-Sectional and Longitudinal Data. *Journal of Econometrics*, 121(1-2): 175–212.

Rauch, James E. 1993. Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. *Journal of Urban Economics*, 34: 380–400.

Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2007. Urban Structure and Growth. *Review of Economic Studies*, 74(2): 597–624.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. Integrated Public Use Microdata Series: Version 5.0. University of Minnesota, Minneapolis, Minn.

Saiz, Albert. Forthcoming. The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics*.

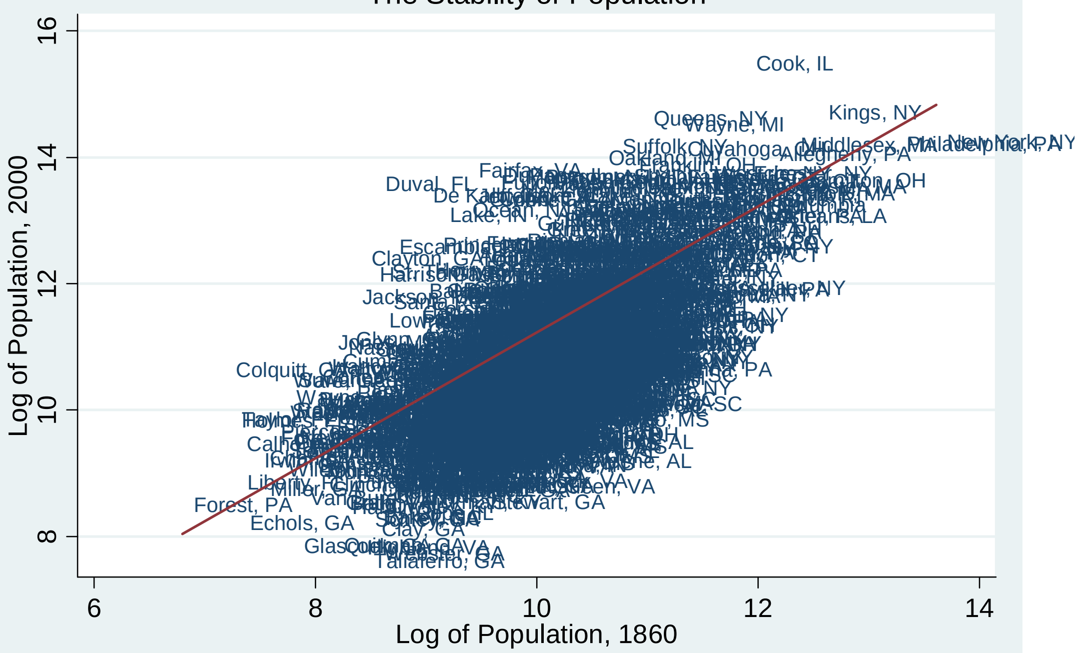
Shapiro, Jesse M. 2006. Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital. *Review of Economics and Statistics*, 88(2): 324–335.

Simon, Curtis J., and Clark Nardinelli. 2002. Human capital and the rise of American cities, 1900–1990. *Regional Science and Urban Economics*, 32: 59–96.

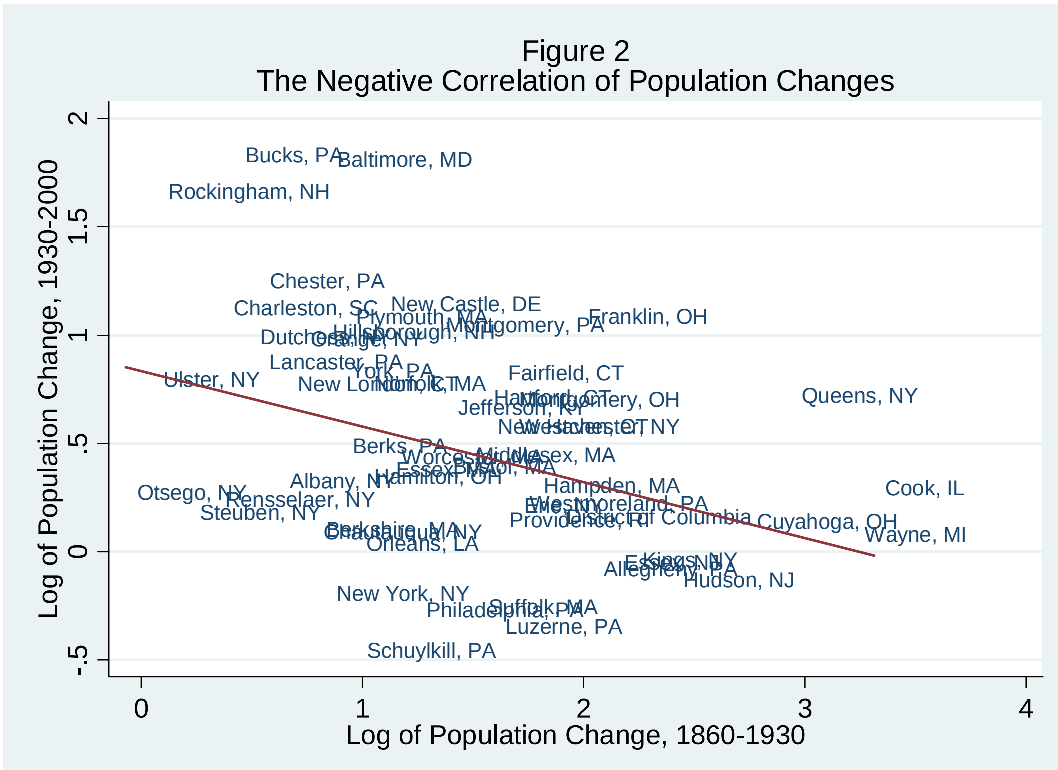
Steckel, Richard H. 1978. The Economics of U.S. Slave and Southern White Fertility. *Journal of Economic History*. 38(1): 289–291.

U.S. Bureau of the Census 1986. County Business Patterns, 1977: U.S. Summary, State, and County Data. Inter-university Consortium for Political and Social Research, Ann Arbor, Mich.

Figure 1
The Stability of Population



Source: County-level U.S. Census data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.



Note: Figure shows the 54 counties that had more than 50,000 people in 1860.
Source: County-level U.S. Census data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

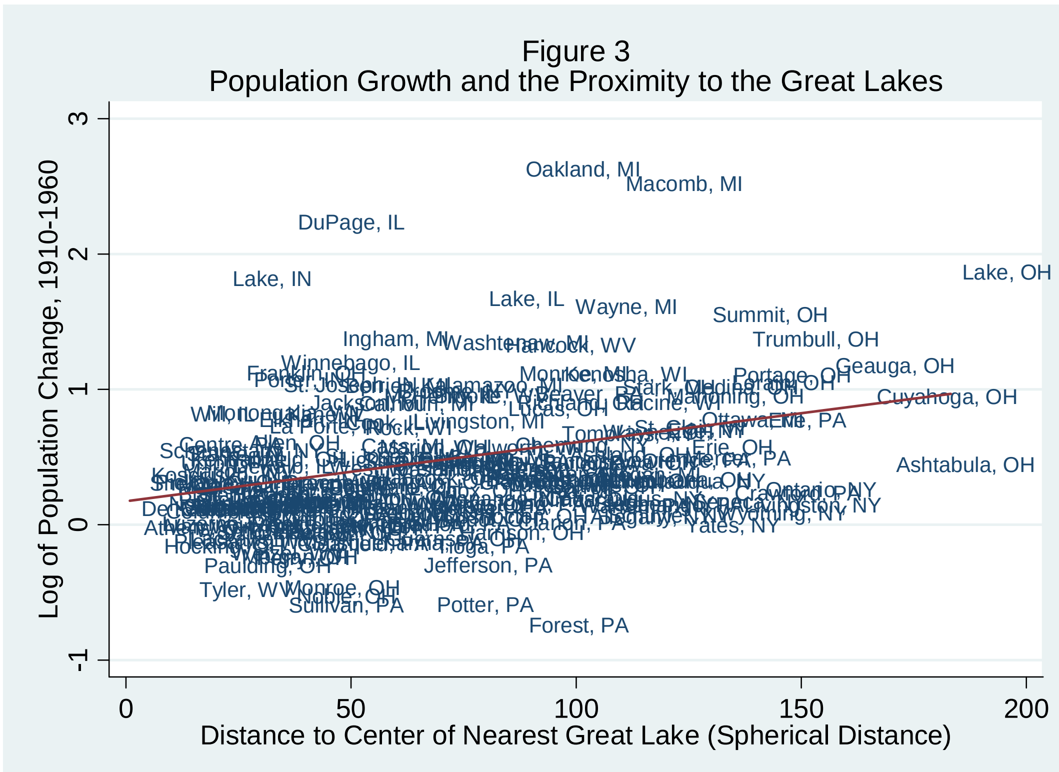
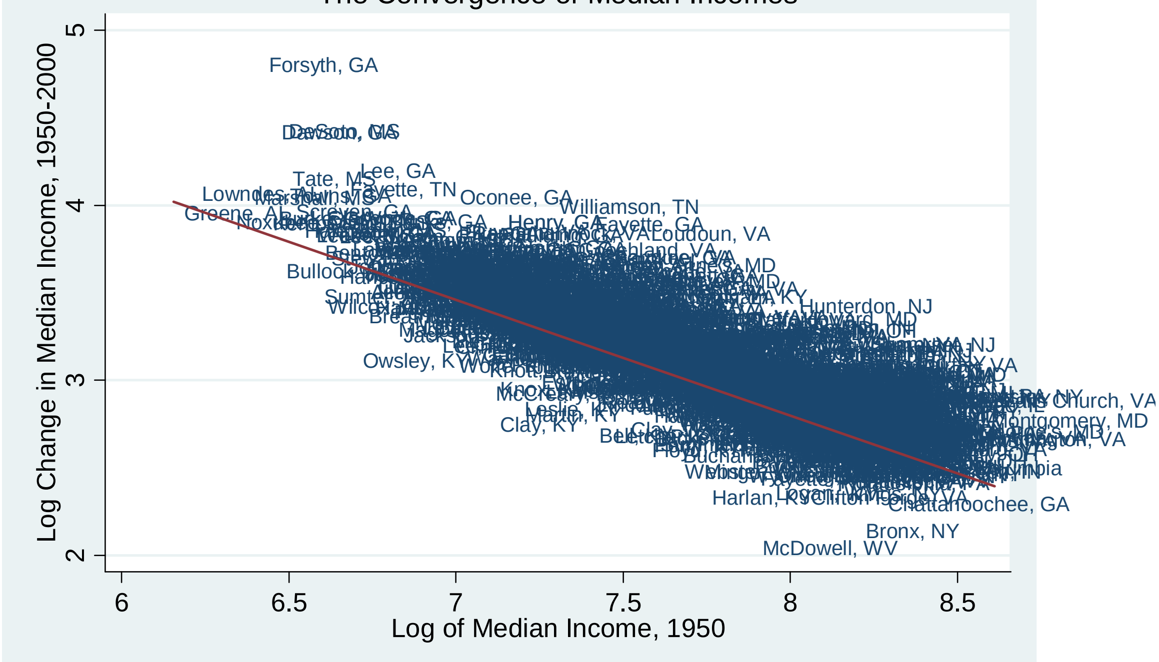


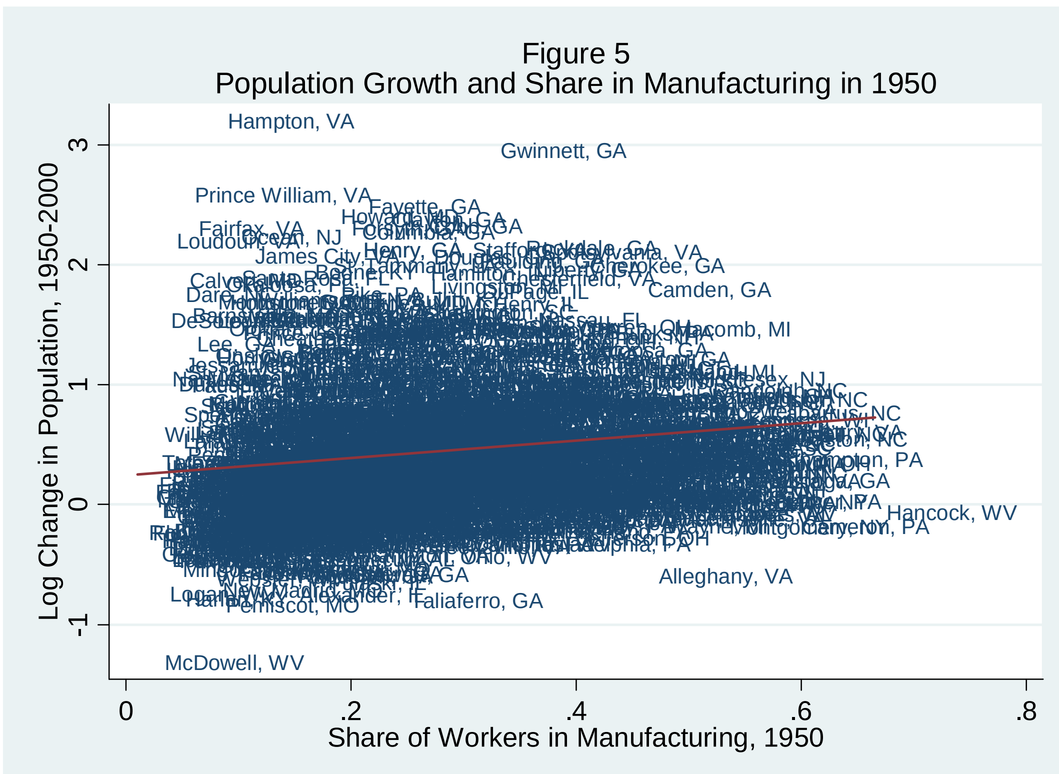
Figure 3
Population Growth and the Proximity to the Great Lakes

Note: Figure shows the counties that are within 200 miles of a Great Lake.
Source: County-level U.S. Census data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

Figure 4
The Convergence of Median Incomes



Source: County-level U.S. Census data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.



Source: County-level U.S. Census data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

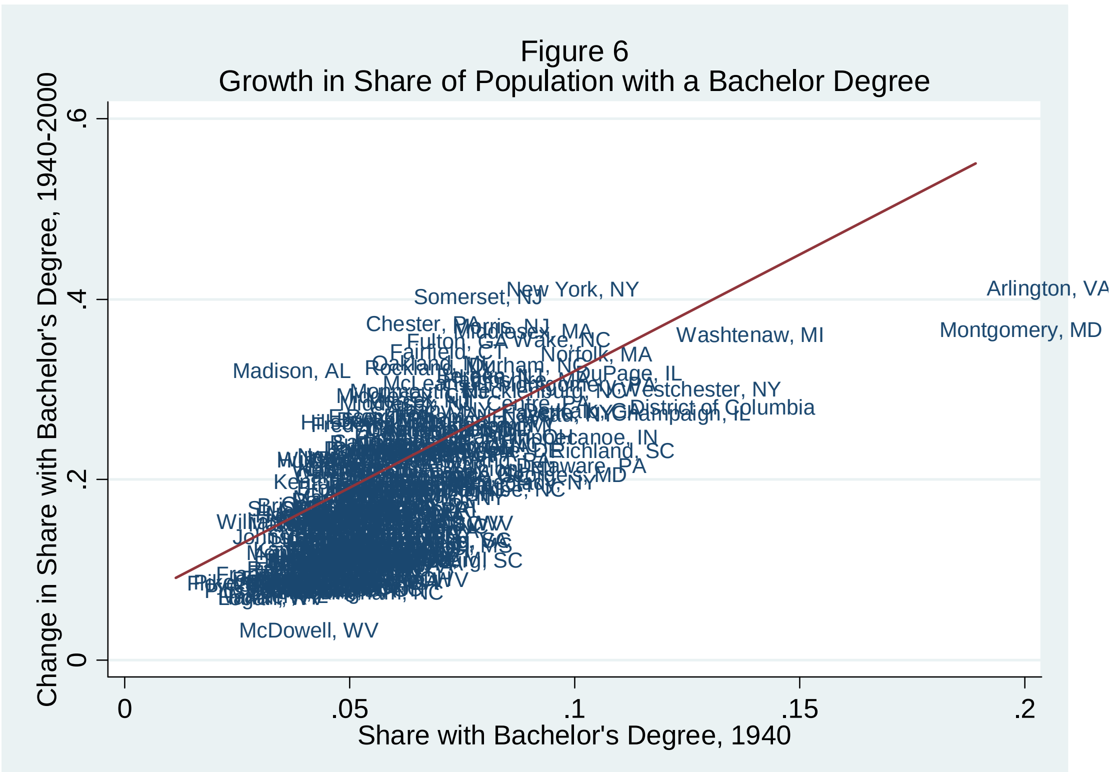
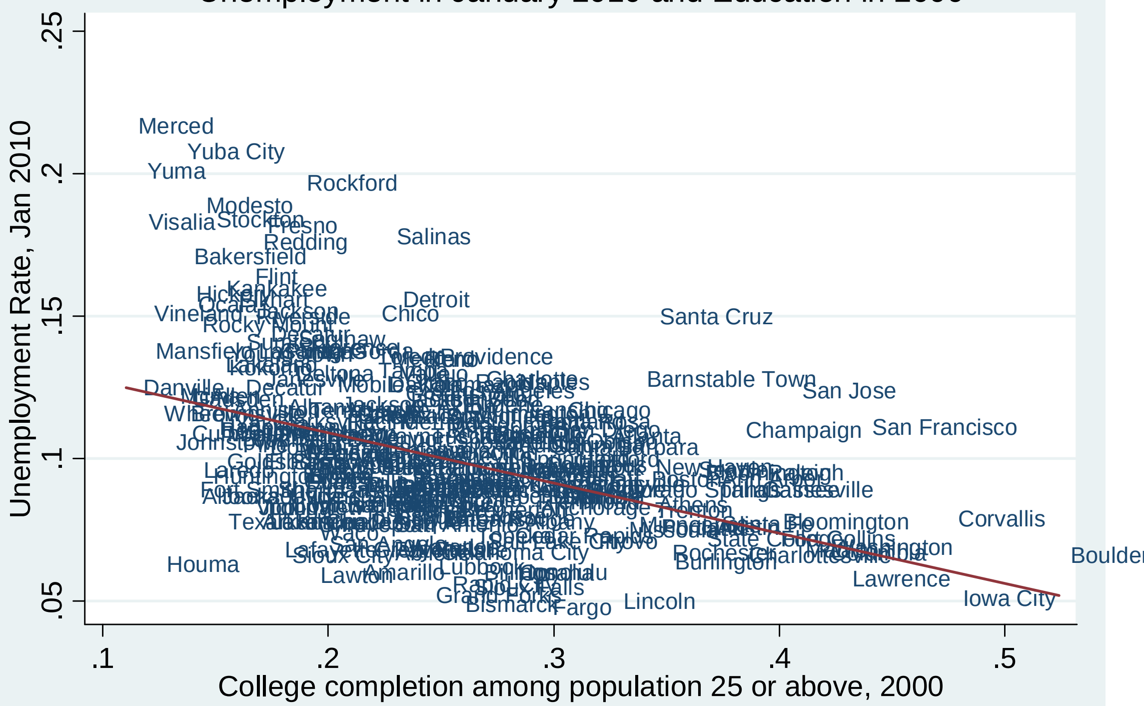


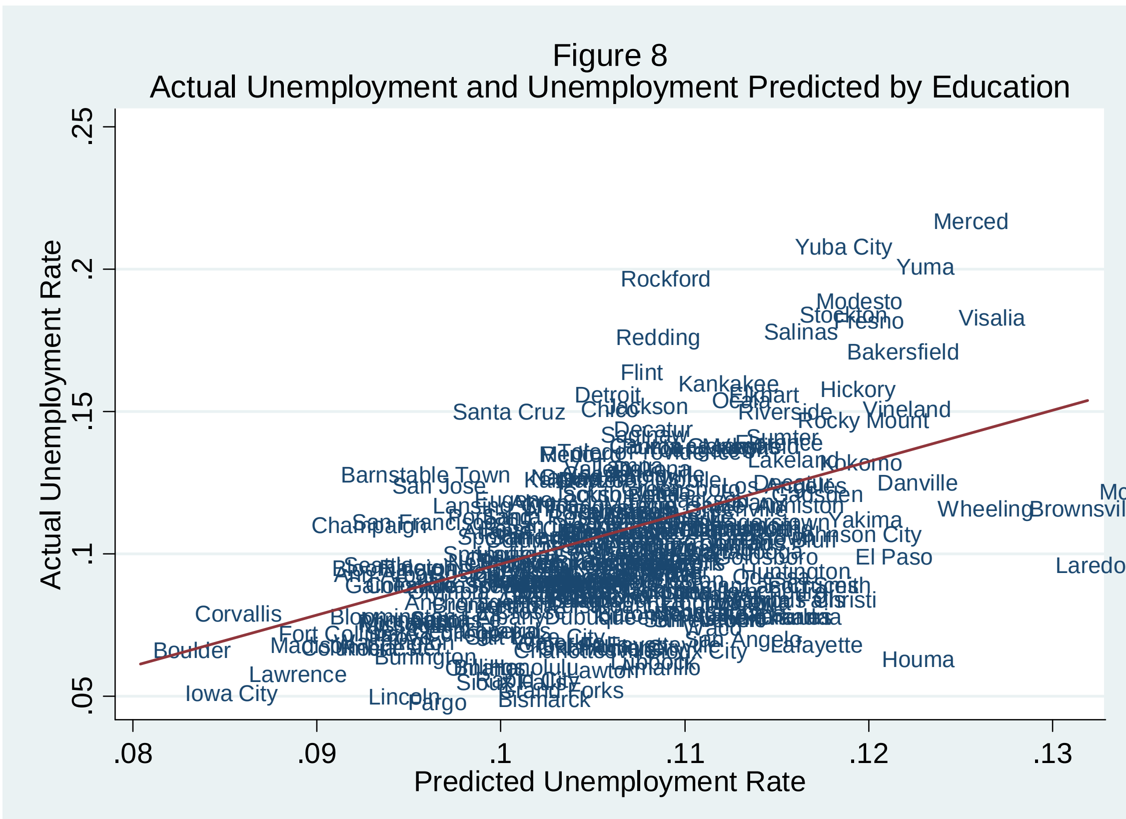
Figure 6
Growth in Share of Population with a Bachelor Degree

Note: Figure shows the counties that had more than 50,000 people in 1940.
Source: County-level U.S. Census data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

Figure 7
Unemployment in January 2010 and Education in 2000



Source: Metropolitan Statistical Area level data from the U.S. Census.



Source: Metropolitan Statistical Area level data from the U.S. Census, and Actual Unemployment Rate from the U.S. Department of Labor, Bureau of Labor Statistics.

Table 1: Population Growth Correlations

Decades	(1) Correlation with Lagged Population Change	(2) Correlation with Lagged Population Change (50,000+)	(3) Correlation with Initial Log Population	(4) Correlation with Initial Log Population (50,000+)
1790s	.	.	-0.4681	-0.9505
1800s	0.3832	0.6462	-0.5625	0.1316
1810s	0.3256	0.4766	-0.5674	-0.0463
1820s	0.4423	0.5231	-0.5136	0.4178
1830s	0.4452	0.9261	-0.6616	0.241
1840s	0.4634	0.8978	-0.5122	0.3922
1850s	0.4715	0.7661	-0.319	-0.0392
1860s	0.3985	0.4631	0.0111	0.0065
1870s	-0.1228	0.4865	-0.3614	-0.0205
1880s	0.3978	0.4541	-0.1252	0.3323
1890s	0.4935	0.5382	-0.1181	0.3691
1900s	0.4149	0.6454	0.1754	0.2947
1910s	0.5027	0.5778	0.2747	0.0903
1920s	0.476	0.4675	0.3381	0.1494
1930s	0.3005	0.4887	0.0415	-0.1585
1940s	0.4151	0.6752	0.3863	-0.0649
1950s	0.7397	0.7327	0.3985	0.0444
1960s	0.7225	0.8196	0.2922	0.0311
1970s	0.3821	0.4349	-0.2247	-0.4462
1980s	0.641	0.7096	0.1062	-0.0693
1990s	0.737	0.7863	-0.0197	-0.157

Source: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

Table 2:
Geography Correlation Tables

	(1)	(2)	(3)
Decades	**Correlation with Longitude**	**Correlation with Proximity to Great Lake**	**Correlation with January Temperature**
1790s	-0.2646	0.3746	-0.0008
1800s	-0.4368	0.4307	-0.226
1810s	-0.3496	0.4473	-0.1891
1820s	-0.2857	0.3053	-0.1514
1830s	-0.3304	0.2631	-0.2676
1840s	-0.3414	0.1442	-0.2424
1850s	-0.3145	0.0703	-0.3466
1860s	-0.1495	0.1028	-0.3229
1870s	-0.046	-0.1188	0.2575
1880s	-0.0256	-0.0336	0.1571
1890s	-0.1145	-0.0771	0.2273
1900s	0.1159	0.0153	0.1339
1910s	0.1448	0.1185	-0.005
1920s	0.1733	0.1182	-0.0802
1930s	-0.0144	-0.0462	0.0379
1940s	0.2431	0.1665	-0.13
1950s	0.2401	0.2075	-0.1843
1960s	0.1313	0.0915	-0.1062
1970s	-0.0435	-0.163	0.2088
1980s	0.1974	-0.1107	0.2243
1990s	-0.0027	-0.1567	0.2702

Sources: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000. Geographical information from ESRI GIS data.

Table 3: Population Growth Regressions

	Change in Population					
	(1) *1800-1830*	(2) *1830-1860*	(3) *1870-1900*	(4) *1900-1930*	(5) *1940-1970*	(6) *1970-2000*
Average January Temperature	-0.025	-0.033	0.008	0.007	-0.002	0.009
	(0.003)**	(0.003)**	(0.001)**	(0.001)**	(0.002)	(0.001)**
Distance to Center of Nearest Great Lake	0.008	0.004	0.001	0.001	0.002	-0.001
	(0.001)**	(0.001)**	(0.000)*	(0.000)**	(0.000)**	(0.000)
Longitude	-0.038	-0.005	-0.000	0.011	0.017	0.008
	(0.005)**	(0.005)	(0.002)	(0.002)**	(0.003)**	(0.002)**
Log of Population, 1800	-0.255					
	(0.025)**					
Log of Population, 1830		-0.551				
		(0.021)**				
Log of Population, 1870			-0.126			
			(0.014)**			
Log of Population, 1900				0.125		
				(0.013)**		
Log of Population, 1940					0.103	
					(0.012)**	
Log of Population, 1970						-0.021
						(0.008)**
Constant	0.628	6.320	1.379	-0.407	0.523	0.872
	(0.57)	(0.505)**	(0.268)**	(0.263)	(0.28)	(0.213)**
Observations	368	788	1210	1276	1324	1338
R-squared	0.63	0.60	0.14	0.11	0.13	0.09

Note: Standard Errors in parenthesis (* significant at 5%; ** significant at 1%).

Sources: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000. Geographical information from ESRI GIS data.

Table 4:
Income Growth Correlations

Decades	(1) Correlation with January Temperature	(2) Correlation with Lagged Income	(3) Correlation with Lagged Income Growth	(4) Correlation with Share Manuf. In 1950
1950s	0.4023	-0.5692		-0.1215
1960s	0.4807	-0.7732	0.2888	-0.4119
1970s	0.3107	-0.6857	0.3303	-0.4911
1980s	0.1842	0.0904	-0.2839	0.086
1990s	0.07	-0.3492	-0.1966	-0.271

Source: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

Table 5:
Education Correlations

Decades	(1) Population Correlation with Lagged BA Share	(2) Income Correlation with Lagged BA Share	(3) Population Correlation with Lagged BA Share (100,000+)	(4) Income Correlation with Lagged BA Share (100,000+)
1940s	0.5904		0.3332	
1950s	0.482	-0.2517	0.3634	0.0291
1960s	0.3758	-0.3864	0.346	0.1586
1970s	-0.0961	-0.369	0.1122	-0.0391
1980s	0.3194	0.3564	0.3908	0.4739
1990s	0.1269	-0.2334	0.2396	-0.1017

Source: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

Table 6: Income and Population Growth Regressions, 1950-2000

	Income Growth	*Population Growth*
Share of Workers in Manufacturing, 1950	0.3025	0.5597
	(0.05)	(0.1369)
Log of Population, 1950	-0.0868	-0.2817
	(0.0139)	(0.0381)
Mean January Temperature	-0.0003	0.0198
	(0.0008)	(0.0022)
Longitude	0.0048	0.0107
	(0.0012)	(0.0032)
Distance to Center of Nearest Great Lake	-0.0009	-0.0007
	(0.0002)	(0.0006)
Share with Bachelor Degrees, 1950	2.5141	4.3104
	(0.3098)	(0.8479)
Log of Population/Bachelor Degree Interaction, 1	1.1749	2.7005
	(0.2127)	(0.5822)
Log of Median Income, 1950	-0.7392	0.4600
	(0.0221)	(0.0605)
Constant	8.8912	-3.2321
	(0.2083)	(0.57)
Observations	1328	1328
R-squared	0.7476	0.1833

Sources: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000. Geographical information from ESRI GIS data.

Table 7: Education and Firm Size Correlations with Population Growth

Decades	(1) Correlation with Share of BAs in 1940	(2) Correlation with Share of BAs in 1940 (50,000+)	(3) Correlation with avg. est. size 1977	(4) Correlation with avg. est. size 1977 (50,000+)
1790s	0.0105	-0.309	0.1152	0.2688
1800s	-0.1012	0.3758	0.0627	0.7698
1810s	-0.096	-0.2574	0.0142	0.391
1820s	-0.0543	0.3583	0.1338	0.7404
1830s	-0.0102	0.5014	0.093	0.7733
1840s	-0.008	0.381	0.113	0.5929
1850s	0.0208	0.1145	0.0651	0.0149
1860s	0.1457	0.0671	0.0779	0.2524
1870s	-0.1386	-0.0157	0.0134	0.2407
1880s	0.0079	0.1089	0.1676	0.3557
1890s	-0.1269	0.0522	0.0751	0.2893
1900s	0.1711	0.2133	0.222	0.2529
1910s	0.2265	0.1866	0.3172	0.3638
1920s	0.4162	0.3581	0.3476	0.2414
1930s	0.2304	0.3216	0.1594	0.0225
1940s	0.5904	0.5613	0.3336	0.1356
1950s	0.4953	0.3619	0.2273	0.0286
1960s	0.383	0.3298	0.1259	-0.0974
1970s	-0.1614	-0.1199	-0.1786	-0.353
1980s	0.1129	0.0806	-0.0862	-0.3212
1990s	-0.0878	-0.1116	-0.1715	-0.2893

Source: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000.

Table 8:
Income and Population Growth Regressions, 1980-2000

	Log Change in Population, 1980-2000		***Log Change in Median Income, 1980-2000***	
	(1)	(2)	(3)	(4)
	Full Sample	*Counties with 50,000+*	*Full Sample*	*Counties with 50,000+*
Share of Workers in Manufacturing, 1980	0.338	0.600	0.390	0.434
	(0.063)**	(0.117)**	(0.031)**	(0.052)**
Log of Population, 1980	-0.017	-0.039	0.001	0.008
	(0.007)*	(0.013)**	(0.003)	(0.006)
Share with Bachelor's Degree, 1980	0.493	0.830	0.966	0.846
	(0.145)**	(0.188)**	(0.071)**	(0.084)**
Distance to Center of Nearest Great Lake	0.000	0.000	0.000	0.000
	(0.000)*	(0.000)**	(0.000)**	(0.000)**
Average Establishment Size, 1977	-0.016	-0.022	-0.011	-0.012
	(0.002)**	(0.003)**	(0.001)**	(0.001)**
Log of Median Income, 1980	0.519	0.646	-0.065	0.062
	(0.039)**	(0.071)**	(0.019)**	(0.032)
Longitude	0.005	0.001	0.006	0.007
	(0.002)**	(0.002)	(0.001)**	(0.001)**
Mean January Temperature	0.010	0.009	-0.003	-0.004
	(0.002)**	(0.002)**	(0.001)**	(0.001)**
Constant	-4.629	-6.027	1.982	0.737
	(0.382)**	(0.663)**	(0.187)**	(0.297)*
Observations	1336	444	1336	444
R-squared	0.28	0.45	0.31	0.52

Note: Standard Errors in parenthesis (* significant at 5%; ** significant at 1%).

Sources: County level data from ICPSR 2896 - Historical, Demographic, Economic, and Social Data: The United States, 1790-2000. Geographical information from ESRI GIS data. Average establishment size in 1977 from County Business Patterns.

Table 9: Metropolitan Area Level Regressions

	Log Change in Population, 1970-2000		*Log Change in Median Income in 2000 $, 1970-2000*		*Log Change in Median Housing Value in 2000 $, 1970-2000*		*Log Change in PopulationAverage Firm Size, 1970-2000*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Population, 1970	-0.007	-0.007	0.003	0.003	0.056	0.057	-0.026	-0.027
	(0.019)	(0.019)	(0.006)	(0.006)	(0.013)**	(0.013)**	(0.013)*	(0.013)*
Log Median Income in 2000 $, 1970	-0.769	-0.841	-0.391	-0.403	-0.297	-0.328	-0.650	-0.643
	(0.191)**	(0.191)**	(0.061)**	(0.062)**	(0.133)*	(0.135)*	(0.131)**	(0.135)**
Log Median Housing Value in 2000 $, 1970	0.273	0.272	0.173	0.174	-0.008	0.004	0.046	0.028
	(0.117)*	(0.115)*	(0.037)**	(0.038)**	(0.081)	(0.082)	(0.080)	(0.082)
South Dummy	0.146	-0.133	0.030	0.015	0.028	0.012	-0.114	-0.163
	(0.054)**	(0.122)	(0.017)	(0.040)	(0.038)	(0.087)	(0.037)**	(0.087)
East Dummy	-0.054	-0.077	-0.010	-0.039	0.054	0.041	-0.164	-0.282
	(0.057)	(0.158)	(0.018)	(0.052)	(0.040)	(0.112)	(0.039)**	(0.112)*
West Dummy	0.384	0.632	-0.044	-0.145	0.299	0.052	0.012	0.084
	(0.051)**	(0.135)**	(0.016)**	(0.044)**	(0.035)**	(0.096)	(0.035)	(0.096)
College completion among population 25 or above, 1970	1.528		0.797		0.802		0.800	
	(0.445)**		(0.142)**		(0.310)*		(0.307)**	
South Dummy * Percent BA in 1970		3.840		0.673		0.405		1.188
		(0.772)**		(0.252)**		(0.548)		(0.548)*
East Dummy * Percent BA in 1970		1.498		0.839		0.424		1.859
		(1.310)		(0.428)		(0.929)		(0.929)*
West Dummy * Percent BA in 1970		-0.573		1.364		2.257		0.211
		(0.812)		(0.265)**		(0.576)**		(0.576)
Midwest Dummy * Percent BA in 1970		1.314		0.583		0.363		0.731
		(0.597)*		(0.195)**		(0.424)		(0.424)
Constant	5.264	6.074	2.275	2.407	2.801	3.040	6.741	6.888
	(1.576)**	(1.595)**	(0.504)**	(0.521)**	(1.100)*	(1.132)**	(1.086)**	(1.132)**
Observations	257	257	257	257	257	257	257	257
R-squared	0.427	0.466	0.379	0.396	0.339	0.362	0.313	0.321

Note: Standard errors in parentheses (* significant at 5%, ** significant at 1%)
Source: Metropolitan Statistical Area data and County Business Patterns from the US Census.

Table 10:
Individual Level Regressions

	Log Real Wage in $ 2000						Log Real Housing Value in $2000	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% of Pop 25+ with a BA (1970), MSA, year 2000 dummy*	0.557	0.472					3.339	
	(0.195)**	(0.121)**					(1.278)**	
*% of Pop 25+ with a BA (1970), Industry, * year 2000 dummy*			0.089					
			(0.016)**					
*% of Pop 25+ with a BA (1970), Industry-MSA, * year 2000 dummy*				0.028	0.002			
				(0.008)**	(0.009)			
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects Detail	MSAs	MSAs and Industry	MSA-Year and Industry	MSA-Year and Industry	MSA-Year and Industry-Year	MSAs	MSAs	MSAs
*% of Pop 25+ with a BA (1970), MSA, * year 2000 dummy * Northeast dummy*						0.788		1.649
						(0.285)**		(0.738)*
*% of Pop 25+ with a BA (1970), MSA, * year 2000 dummy * Midwest dummy*						0.599		0.435
						(0.289)*		(0.690)
*% of Pop 25+ with a BA (1970), MSA, * year 2000 dummy * South dummy*						0.631		1.246
						(0.188)**		(0.632)*
*% of Pop 25+ with a BA (1970), MSA, * year 2000 dummy * West dummy*						0.589		3.734
						(0.224)**		(0.664)**
Constant	10.400	7.254	7.952	7.426	8.783	10.634	10.843	10.853
	(0.017)**	(0.286)**	(0.307)**	(0.313)**	(0.344)**	(0.019)**	(0.063)**	(0.062)**
Observations	402490	402490	377891	364266	364266	402490	426295	426295
R-squared	0.33	0.38	0.38	0.38	0.39	0.33	0.36	0.36

Notes:
(1) Robust standard errors in parentheses. Standard errors clustered by MSA, Industry, and Year (1)-(5) or MSA-Year (6)-(8)
(2) Wage regressions data only for males 25-55, who are in the labor force, who worked 35 or more hours per week and 40 or more weeks per year, and who earned over a certain salary (equal or more than if they had worked half-time at minimum wage).
(3) Individual-level data (wages, housing prices, and controls) from IPUMS
(4) MSA-level data (baPct1970 in MSA, Population, Log of Real Housing Value, and Log of Real Income) from aggregate level Census data.
(5) Wage regressions includes controls for individual education, age, and race, as well as the interaction of those variables with a dummy variable for year 2000. Housing regressions include housing quality controls as well as the interaction of those variables with a dummy variable for year 2000.
(6) Industry and Industry-MSA BA Shares calculated using IPUMS data.
(7) Includes controls for initial (1970) values for population, median income, and median housing value interacted with a year 2000 dummy.

Table 11:
Estimated Coefficients

	MSA-level Coefficients					*Individual-level Coefficients*				
	$\tilde{\beta}_s^E$	β_s^θ	$\beta_s^{\bar{L}}$	β_s^a	β_s^E	$\tilde{\beta}_s^E$	β_s^θ	$\beta_s^{\bar{L}}$	β_s^a	β_s^E
Nation	4.716	-0.556	1.523	1.329	0.729	3.757	0.444	-1.253	1.009	0.729
	(0.853)	(0.111)	(0.517)	(0.249)	(0.573)	(0.282)	(0.061)	(0.085)	(0.096)	(0.573)
East	4.855	-0.712	1.913	1.739	-0.361	4.65	-0.293	0.638	1.728	-0.361
	(2.551)	(0.333)	(1.312)	(0.701)	(1.212)	(0.409)	(0.116)	(0.166)	(0.136)	(1.212)
Midwest	3.647	-0.475	1.534	1.022	0.582	3.711	-0.469	1.478	1.092	0.582
	(1.027)	(0.143)	(0.462)	(0.317)	(0.747)	(0.337)	(0.079)	(0.095)	(0.112)	(0.747)
South	6.531	-0.551	4.108	1.293	2.652	6.366	-0.258	3.225	1.276	2.652
	(1.702)	(0.207)	(0.956)	(0.461)	(0.811)	(0.304)	(0.075)	(0.091)	(0.099)	(0.811)
West	4.883	-0.687	-1.466	1.889	-0.784	1.785	0.531	-3.717	0.91	-0.784
	(1.846)	(0.215)	(1.114)	(0.531)	(1.22)	(0.341)	(0.099)	(0.135)	(0.117)	(1.22)

Notes:

(1) MSA-level coefficients are from Table 9. Individual-level coefficients are from Table 10.

(2) Values used were σ=4 and μ=.7. See section III for formulas.