



The divergence of human capital levels across cities*

Christopher R. Berry¹, Edward L. Glaeser²

¹ The Harris School, The University of Chicago, 1155 E. 60th St., Chicago, IL 60637, USA
(e-mail: crberry@uchicago.edu)

² Department of Economics, Harvard University and NBER, John F. Kennedy School of Government, 79 JFK St., Cambridge, MA 02138, USA (e-mail: eglaeser@harvard.edu)

Received: 18 May 2005 / Accepted: 28 August 2005

Abstract. Over the past 30 years, the share of adult populations with college degrees increased more in cities with higher initial schooling levels than in initially less educated places. This tendency appears to be driven by shifts in labor demand as there is an increasing wage premium for skilled people working in skilled cities. In this article, we present a model where the clustering of skilled people in metropolitan areas is driven by the tendency of skilled entrepreneurs to innovate in ways that employ other skilled people and by the elasticity of housing supply.

JEL classification: J24, R12, R23, R31

Key words: Metropolitan, urban, city, human capital

1 Introduction

U.S. metropolitan areas with more college graduates in 1990 became increasingly skilled over the 1990s. Figure 1 shows a 52 percent correlation between the initial share of adults with college degrees in 1990 and the growth in the share of adults with college degrees between 1990 and 2000 across metropolitan areas. A one percent increase in the share of adults with college degrees in the 1990s is associated with a 0.13 percent increase in the share of the population with college degrees between 1990 and 2000. The correlation between initial share of the population with college degrees and growth in that share was 44 percent in the 1980s and 60 percent in the 1970s.

* Glaeser thanks the Taubman Center for State and Local Government for financial support. Lawrence Katz and three anonymous referees provided very helpful comments.

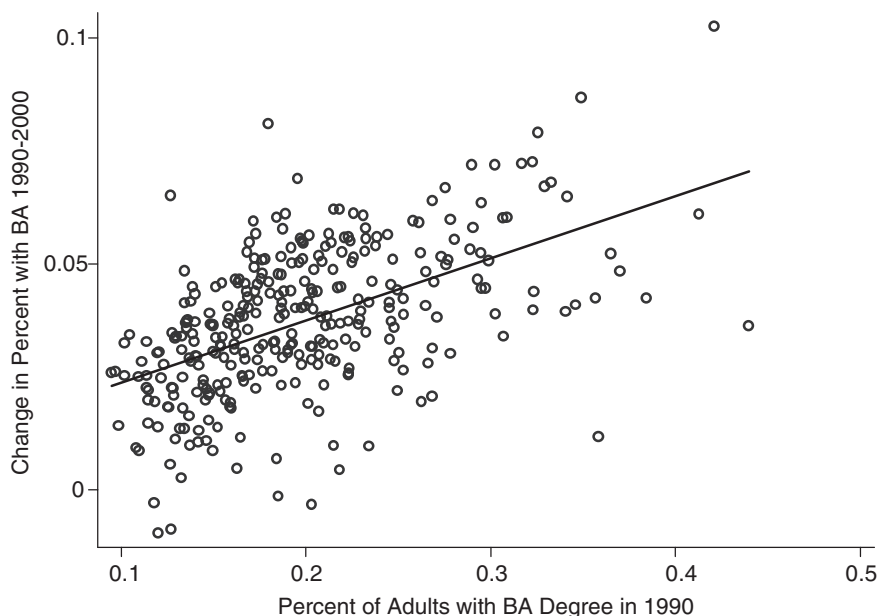


Fig. 1. Initial attainment and attainment growth, 1990–2000

Note for Figures 1 and 2: Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England), using constant 1999 boundaries. Data are from the census.

In Section 2 of this article, we document this divergence of skill levels. There is a strong correlation between changes in the share of the adult population with college degrees and the initial share of the population that is well educated and this is robust to a wide number of controls. This relationship is also robust to following Moretti (2004) and using colleges per capita in 1940 as an instrument for initial skill levels. This tendency of initially skilled places to become more skilled over time has, so far, caused only very modest increases in segregation by skill across metropolitan areas. Traditional measures of segregation such as isolation and dissimilarity indices show a small rise over the past 30 years. Still, segregation by skill across metropolitan area remains quite modest.

In Section 3, we present a simple model of urban agglomeration that can potentially explain the increased tendency of skilled people to move to initially skilled areas. The core assumption of the model is that the number of entrepreneurs is a function of the number of skilled (and unskilled) people working in an area. This model departs from traditional regional models by assuming that new entrepreneurs are, at least for a time, relatively immobile. If skilled people are more likely to innovate in ways that employ other skilled people then this creates an agglomeration economy where skilled people want to be around each other.

The model can explain the observed changes over time through two mechanisms. First, if the tendency of skilled entrepreneurs to particularly hire skilled people has increased over time, then this would explain why skilled people come to initially skilled cities. Second, if housing supply has become more inelastic

over time, then unskilled people will face increasingly higher prices to live around skilled people and move to cheaper places.

In Section 4, we evaluate these explanations. Cross-industry evidence shows that there has been more innovation in skilled sectors of the economy and new firms generally employ the skilled (Abowd et al. 2001). We find an increasing correlation across industries between managerial skills and worker skills. In 1970, the correlation across industries between the share of managers with college degrees and the share of workers with college degrees was 38 percent. In 2000, that same correlation coefficient reached 51 percent. A one percentage point increase in the share of managers with college degrees in 1970 was associated with a 0.15 percentage point increase in the share of workers with college degrees. In 2000, a one percentage point increase in the share of managers with college degrees was associated with a 0.38 percentage point increase in share of workers' college degrees. The tendency of skilled managers to employ skilled workers has clearly increased over time. We take all of this as evidence supporting the view that skilled entrepreneurs increasingly provide employment opportunities primarily for skilled workers.

The evidence for the housing supply inelasticity hypothesis is more mixed. While there is significant evidence suggesting that housing supply is more inelastic now than in the past, there is little evidence suggesting that this explains the increased tendency of skilled people to move into skilled areas. The relationship between metropolitan area skill levels and housing prices has remained relatively constant between 1980 and today. Controlling for initial housing price levels does not reduce the correlation between initial skills and later skill growth.

In Section 5 of this article, we explore the predictions of the model for wage and income patterns across metropolitan areas. As the model suggests, there is an increasingly tight relationship between education and income at the metropolitan area level. In 1970, the correlation between share of adults with college degrees and the log of income in the metropolitan area was 21 percent. In 2000, that correlation rose to 63 percent. When we hold individual education constant, the correlation between metropolitan area education and individual wages has also risen and this rise has been much more pronounced for high skilled workers. This finding supports the view that our core fact is the result of increasing demand for high skilled workers in initially high skill cities.

Finally, we document that the divergence of skill levels seems related to the decline of income convergence across metropolitan areas (Barro and Sala-i-Martin 1992). In the 1970s, the correlation between initial wages and wage growth was -36 percent. In the 1990s, this correlation switched signs and became 15 percent. While the end of regional convergence certainly does not imply a need for policy intervention, it does at least mean that we cannot expect that differences in income across space will naturally disappear.

2 Evidence on skill divergence

In this section, we review the basic facts that motivate this article. First, we present the evidence on the relationship between changes in the share of skilled

individuals at the metropolitan area level and the initial share of individuals who are skilled. Second, we examine whether the increase in skills can be explained by industrial shifts in skilled cities. Third, we present evidence on the segregation of the skilled within the United States.

2.1 *The correlation between increase in skills and initial skills*

We begin with our most basic relationship: the connection between growth in the percentage of adults with college degrees and the initial share of people in the area with college degrees. We estimate:

$$College_{t+1} - College_t = \alpha + \beta \cdot College_t + Other\ Controls + \varepsilon, \quad (1)$$

where $College_t$ refers to the share of the adult population at time t with sixteen years of schooling or more. Our units of observation are all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England). Unless otherwise noted, all data come from the 1970, 1980, 1990, and 2000 US censuses. We rely on county level data from each census aggregated to a consistent set of 1999 metropolitan area boundaries.¹

Table 1 shows our first set of regressions. Regression (1) of Table 1 quantifies the relationship, shown in Figure 1, between the growth in the share of population with college degrees and the initial share of the population with degrees in the 1990s with no other controls. In this regression, the R^2 is 28 percent and the coefficient is 0.14. As the share of college graduates in the metropolitan area increased by 1 percent in 1990, on average the share of adults with college degrees increased by 1.14 percentage points in 2000.

This coefficient does not merely reflect a tendency of skilled people to come to certain areas in recent years. Following Moretti (2004), we can use the number of colleges per capita in the metropolitan area in 1940 as a measure of the area's long term commitment to human capital. There is a 23 percent correlation between the number of colleges in the metropolitan area in 1940 and the growth in the share of adults with college degrees between 1990 and 2000. When we use the number of colleges per capita in 1940 as an instrument for the share of the adult population with college degrees in 1990, we estimate:

$$College_{2000} - College_{1990} = -0.006 + 0.217 \cdot College_{1990}. \quad (2)$$

(0.009) (0.048)

The relationship between initial skills and growth in skill is, if anything, stronger, when we use this historic measure of human capital as an instrument for schooling in 1990.

In regression (2) of Table 1, we include standard controls for four region dummies, the logarithm of the initial level of income, the logarithm of the initial population level, and the initial share of the population working in manufactur-

¹ The data set is described further in the appendix of Glaeser and Saiz (2004).

Table 1. Human capital divergence across MSAs

	Change in percentage of adults with bachelor's degree		Change in percentage of adults with bachelor's degree		Change in percentage of adults with bachelor's degree	
	1990–2000		1980–1990		1970–1980	
	(1)	(2)	(3)	(4)	(5)	(6)
Share with bachelor's degree (age 25+) at $t - 10$	0.138 (0.0124)	0.117 (0.014)	0.130 (0.0147)	0.147 (0.0137)	0.243 (0.018)	0.205 (0.0199)
Log of population at $t - 10$		0.0005 (0.0008)		0.005 (0.00077)		0.003 (0.0008)
Share of workers in manufacturing at $t - 10$		0.0279 0.0117		0.01837 (0.009)		0.0222 (0.009)
Log of per capita income at $t - 10$		0.025 (0.006)		0.005 (0.005)		0.0117 (0.006)
Regional fixed effects	No	Yes	No	Yes	No	Yes
Observations	318	318	318	318	318	318
R^2	0.28	0.40	0.20	0.52	0.36	0.45

Notes: Standard errors in parentheses. Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England). All data are from the 1970, 1980, 1990 and 2000 US censuses. County level data from each census are aggregated to a consistent set of metropolitan area boundaries as defined by the Office of Management and Budget in 1999.

ing industries. Regional effects are weak and only the west had a discernibly lower level of skill upgrading. There is also no correlation between city size and growth of skills. Manufacturing cities had a slightly higher increase in the share of their populations with college degrees. Initial income also predicts an increase in the share of the growth in college graduates. While many of these effects are significant, they only modestly reduce the coefficient on the initial share of the population with college degrees from 0.14 to 0.12.

Regression (3) of Table 1 reproduces regression (1) for the 1980s. In this decade the correlation between initial skills and subsequent skill growth is weaker, but it is still significant. The correlation coefficient between growth and initial level is 44 percent. The coefficient in this regression is 0.13 which is almost the same as in the 1990s. Figure 2 shows this relationship. Comparing Figures 1 and 2 shows that in the 1990s, the most highly skilled cities (often based around universities) increased their skill levels substantially, but in the 1980s, the skill growth in these university towns was considerably more modest.

Regression (4) of Table 1 reproduces regression (2) for the 1980s. In this case, including the other controls causes the coefficient on the initial years of schooling to increase. One change between the 1980s and 1990s is that initial income does not predict an increase in the share of the adult population with college degrees between 1980 and 1990. A second change is that metropolitan areas with more people became more skilled in the 1980s.

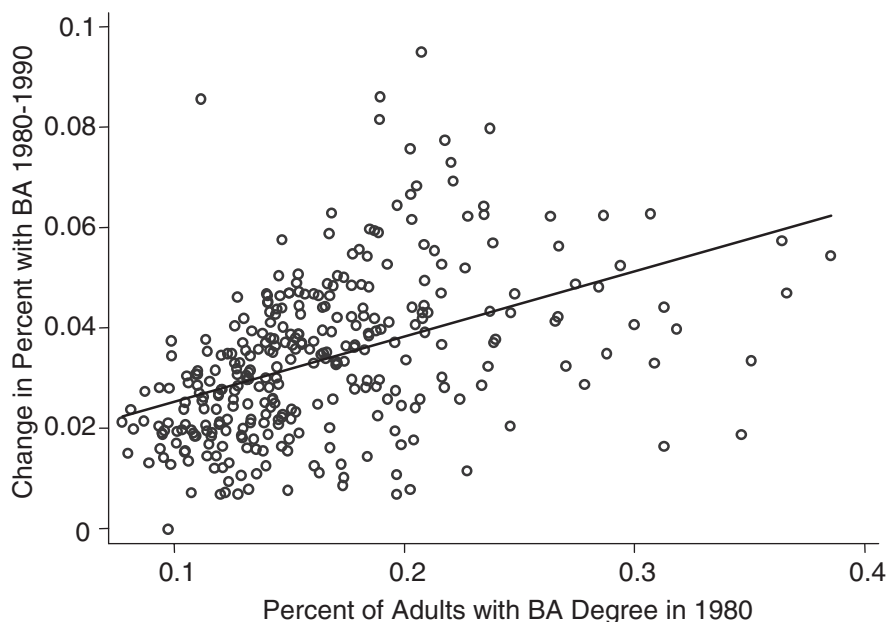


Fig. 2. Initial attainment and attainment growth, 1980–1990

Regressions (5) and (6) of Table 1 reproduce regressions (1) and (2) for the 1970s. In this decade, the raw correlation between the initial share of the adult population with college degrees and later growth in this measure is at its strongest. The coefficient in regression (5) is 0.24 and the R^2 is 36 percent. Adding other controls again makes little difference to this coefficient. In this decade, both initial population and initial income predict an increase in the share of the population with college degrees, but during this decade manufacturing cities were less likely to increase the share of their population with sixteen years of schooling or more.

The functional form of equation (1) is somewhat problematic since the change in a share cannot be distributed normally. Furthermore, this dependent variable corresponds weakly to the standard structure of urban growth regressions. In Table 2, we instead estimate:

$$\text{Log}\left(\frac{\text{Adults with BAS}_{t+1}}{\text{Adults with BAS}_t}\right) = \alpha_1 + \beta_1 \cdot \text{College}_t + \varepsilon \quad (3)$$

and

$$\text{Log}\left(\frac{\text{Adults w.o. BAS}_{t+1}}{\text{Adults w.o. BAS}_t}\right) = \alpha_2 + \beta_2 \cdot \text{College}_t + \varepsilon \quad (4)$$

Table 2. Initial skills and change in high and low skill populations

	1990–2000		1980–1990		1970–1980	
	Log change in number of adults with college degree	Log change in number of adults w.o. college degree	Log change in number of adults with college degree	Log change in number of adults w.o. college degree	Log change in number of adults with college degree	Log change in number of adults w.o. college degree
	(1)	(2)	(3)	(4)	(5)	(6)
Share with bachelor's degree (age 25+) at $t - 10$	0.192 (0.102)	0.067 (0.0824)	0.312 (0.139)	0.306 (0.126)	0.055 (0.206)	0.596 (0.196)
Observations	318	318	318	318	318	318
R^2	0.011	0.002	0.0157	0.0183	0.0002	0.028
Constant	0.277	0.0794	0.334	0.105	0.643	0.127
Test: coefficient for adults with college degree equals coefficient for adults w.o. college degree						
χ^2 statistic	2.95			0.00		18.93
p -value	0.09			0.95		0.00

Notes: Standard errors in parentheses. Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England). All data are from the 1970, 1980, 1990 and 2000 US censuses. County level data from each census are aggregated to a consistent set of metropolitan area boundaries as defined by the Office of Management and Budget in 1999.

where Adults w BAs, refers to the number of college educated adults in the metropolitan area as of time t and Adults w.o. BAs, refers to the number of adults with less than sixteen years of schooling in the metropolitan area during the same year. We focus on the difference between β_1 , the impact of the initial college variable on growth in the number of adults with sixteen years of schooling or more and β_2 , the impact of the initial college variable on growth in the number of adults with less than sixteen years of schooling. We omit other controls to make comparisons between the coefficients more straightforward.

Regression (1) of Table 2 shows the results for the 1990s. The coefficient on the lagged share of the population with college degrees is positive and significant at the 10 percent level. A 0.01 increase in the share of people with sixteen years of schooling is associated with a 0.002 log point increase in the growth of people with sixteen years of schooling. The coefficient becomes significant at the five percent level and increases by more than 10 percent if Las Vegas is omitted from the sample. As has been previously noted (Glaeser et al. 1995), the correlation between education and growth is often sensitive to the inclusion of that single metropolitan area.

In regression (2) of Table 2, we look at the correlation between the initial share of the population with college degrees and the growth in the adult population without college degrees. In the 1990s, there is essentially no correlation. We are able to reject the equality of the coefficients in regressions (1) and (2) at the 10 percent level. In regressions (3) and (4) of Table 2, we repeat regressions (1) and (2) of Table 2 for the 1980s. In the 1980s, the impact of initial years of schooling on later population growth is essentially identical for both college graduates and non-college graduates. A 0.01 increase in the share of the population with college degrees is associated with 0.003 log point increase in population growth for either population subgroup.

In regressions (5) and (6) of Table 2, we repeat this exercise for the 1970s and find that in that decade the impact of college education is completely reversed. Places that were better educated in 1970 added significant numbers of non-college graduates, but added almost no more college graduates. A 0.01 increase in the share of the population with college degrees is associated with 0.006 log point increase in population growth for non-college graduates but a 0.0006 log point increase in population growth for college graduates. The test of the equality of the coefficients in regressions (5) and (6) soundly rejects.

Table 2 shows the transformation of the impact of college graduates on population growth. In the 1970s, skilled cities grew by attracting unskilled workers. In the 1990s, skilled cities grew by attracting skilled workers. The relationships are generally less striking when we look at the logarithm of population because this concave transformation emphasizes the growth in percent college educated of cities with a small initial number of college graduates, but nonetheless these regressions continue to emphasize a changing pattern where skills are increasingly attracting skills.

Is the connection between initial skills and skill upgrading the result of skilled industries increasingly moving to skilled cities? To test this hypothesis, we construct an index that predicts the level of skills in a metropolitan area based on its

industrial composition and the skill levels of industries at the national level. Each metropolitan area's industrial skill index is:

$$\sum_{\text{industries}} \frac{\text{Employment in Industry in MSA}}{\text{Total Employment in MSA}} \cdot \frac{\text{Workers in Industry w. BAs in U.S.A.}}{\text{Total Workers in Industry in U.S.A.}} \quad (5)$$

The index sums across industries the multiple of the industry's share of MSA employment and the average share of the industry with college degrees in the U.S. as a whole. The index represents the share of workers with BAs in the MSA that would be found if each industry in the MSA had the same share of BA workers as in the industry nationwide. In other words, the index is a prediction of local educational attainment based solely on the MSA's industrial composition. This index is calculated for one-digit industries. The industry share of MSA employment is from the Regional Economic Information System (REIS) of the Bureau of Economic Analysis, and we calculate education levels in each industry using the public use micro-samples (PUMS) of the U.S. Census (Ruggles et al. 2004).

Table 3 reports results from regressions of changes in this index on the initial share of adults in the metropolitan area with sixteen years of schooling or more. In the upper panel we report coefficients from these regressions for the 1970s, 1980s and 1990s. The upper panel shows the effect of initial schooling on increases in the skill level of industries at the metropolitan area level. The first column gives results for the 1970s. In that decade, initial skills did predict a transition into more skilled industries, but this effect is weak relative to the overall connection between initial skills and skills growth. In the 1980s and the 1990s, there is no connection between initial skills and industrial shifts into more skilled industries. While these results are coarse and use only one-digit industries, they certainly suggest that the impact of skills on skill upgrading is occurring within industries. There is at best a small impact of initial skills on shifts into more skilled industries.

In the lower panel of Table 3 our dependent variable is the change over the decade in the difference between the actual share of adults in the metropolitan area with BAs and the industrial skill index. The lower panel shows the impact of initial schooling on increases in the skill level at the metropolitan area level *beyond* the skill changes predicted by changes in industrial composition. In all three decades, the relationship is significant and positive, indicating that initial attainment increases skills primarily through channels other than industrial composition.

2.2 A rise in the segregation of the skilled?

The inevitable counterpart of the correlation between growth in percentage of college educated and the initial share of the population with sixteen years of schooling is that cities are becoming more different in their educational bases. While segregation by skill across metropolitan areas remains extremely modest, we have gone from a situation where skills were remarkably evenly distributed

Table 3. Skills and industrial composition

	1970s	1980s	1990s
Dependent Variable: Change in Industrial Skill Index			
Initial share of BAs	0.035 (0.009)	-0.004 (0.008)	0.003 (0.008)
Observations	318	318	318
R^2	0.049	0.0008	0.0004
Constant	0.070	0.031	0.032
Dependent Variable: Change in MSA Education minus Industrial Skill Index			
Initial share of BAs	0.209 (0.018)	0.134 (0.016)	0.134 (0.014)
Observations	318	318	318
R^2	0.288	0.182	0.219
Constant	-0.045	-0.019	-0.022

Notes: Standard errors in parentheses. Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England). Share of BAs is obtained from the US census. County level data from each census are aggregated to a consistent set of metropolitan area boundaries as defined by the Office of Management and Budget in 1999. The Industrial Skill Index is based on data from the Regional Economic Information System (REIS) of the Bureau of Economic Analysis and the public use micro-samples (PUMS) of the U.S. Census, as defined in the text.

across space to a situation where metropolitan areas increasingly differ from one another on the basis of their human capital levels.

Table 4 shows the time path for education heterogeneity across metropolitan areas. The first column of the table shows the remarkable increase in the mean level of schooling over this period. In 1970, in the average metropolitan area, 11.2 percent of adults had college degrees. Even more remarkable, metropolitan areas were relatively homogeneous in their years of schooling. The standard deviation of the share of the adult population with college degrees in 1970 was 0.042. One half of metropolitan areas had between 8.7 percent college graduates and 13.1 percent college graduates. This is shown in the third column which gives the gap in percent college educated between the metropolitan area that is more educated than exactly 25 percent of all metropolitan areas and the metropolitan area that is more educated than exactly 75 percent of all metropolitan areas.

In 1980, the average share of metropolitan areas with college degrees increased to 16.4 percent and the standard deviation rose to 0.054. The percentage increase in the mean during this period was much larger than the percentage increase in the standard deviation, so the coefficient of variation actually fell. Still, this does not lessen the fact that cities were increasingly differentiating themselves on the basis of years of schooling during this decade. The gap between 25th percentile and 75th percentile in the share of adults with college degrees increased to 6.3 percent.

In the 1990s, the mean share of the population with college degrees increased further to 18.7 percent, and the standard deviation of this number across metro-

Table 4. Segregation by skill

Year	Mean share BAs across Metropolitan Areas	Standard Deviation	25–75 Difference	Dissimilarity Index	Isolation Index
1970	0.112	0.042	0.045	0.110	0.008
1980	0.164	0.054	0.063	0.114	0.011
1990	0.187	0.063	0.077	0.123	0.013
2000	0.226	0.073	0.096	0.128	0.016

Notes: Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England). All data are from the 1970, 1980, 1990 and 2000 US censuses. County level data from each census are aggregated to a consistent set of metropolitan area boundaries as defined by the Office of Management and Budget in 1999. Formulas for the dissimilarity and isolation indices are given in the text.

politan areas increased to 0.063. In 1990, the mean share of metropolitan area population with college degrees increase to 22.6 percent and the standard deviation increased to 0.073. The gap between the 25th percent and the 75th percentile continued to widen to 9.6 percentage points. The rising standard deviation reflects the increasing weight in the tails of the distribution. By 2000, there were 62 metropolitan areas with less than seventeen percent of their adults had college degrees and 32 where more than thirty-four percent of their adults had college degrees. Since the distribution is truncated below at zero, the most impressive growth has been at the upper tail of the distribution. In 1970, no metropolitan area had more than 30.8 percent of its adult population with college degrees. In 2000, there were 49 metropolitan areas where more than 30.8 percent of adult population had sixteen years of schooling or more.

While there is clearly a correlation between change in the share of adults with college degrees and the initial share of the population that is well educated, and it is clear that the variance of these shares is rising over time, the changes in the segregation of the skilled across metropolitan areas remain modest. Also in Table 4, we use standard measures from the segregation literature to assess the degree to which the skilled are segregated across metropolitan areas. Our first measure is the dissimilarity index:

$$Dissimilarity = \frac{1}{2} \sum_{MSAs} \left| \frac{\text{Adults w. } BA_{MSA}}{\text{Total Adults w. BAs}} - \frac{\text{Adults w.o. } BA_{MSA}}{\text{Total Adults w.o. BAs}} \right| \quad (6)$$

where Adults w. BA_{MSA} refers to the number of adults in the metropolitan area with college degrees, Adults w.o. BA_{MSA} refers to the number of adults in the metropolitan area without sixteen years of schooling. This measure can be interpreted as the share of people with sixteen years of schooling who would have to move for there to be a completely even distribution of people with sixteen years of schooling across all metropolitan areas.

Table 4 shows the time path of this index. Dissimilarity does rise over this period from 0.11 to 0.128. These numbers are quite low relative to the segregation of other groups such as African-Americans and they reflect the fact that skills

are still distributed relatively evenly over space. So while skilled people are becoming more segregated, this trend is mild and the skilled still remain relatively integrated over space.

The second primary measure of segregation is the isolation index or:

$$Isolation = \sum_{MSAs} \frac{Adults \text{ w. } BAs_{MSA}}{Adults_{MSA}} \frac{Adults \text{ w. } BAs_{MSA}}{Total \text{ Adults w. } BAs} - \frac{Total \text{ Adults w. } BAs}{Total \text{ Adults}} \quad (7)$$

where $Adults_{MSA}$ refers to the number of adults in the MSA and $Total \text{ Adults}$ refers to the total number of adults across all metropolitan areas. The first term in the expression averages across adults with sixteen years of schooling the average share of adults with sixteen years of schooling in their metropolitan area. The second term subtracts off the average share of adults with sixteen years of schooling in the sample as a whole. This index captures the extent to which people with college degrees are surrounded only by like people.

If we didn't correct for the overall increase in the share of the population with college degrees then the isolation index would rise from 12.8 percent to 28.2 percent over the time period. At the start of the time period, on average someone with a college degree lived in a metropolitan area where about one-eighth of the population had a similar credential. Thirty years later, a typical person with a college degree lives in a metropolitan area where more than one-in-four have the same amount of schooling.

But most of this gain reflects the rise in education generally and when we correct by subtracting the sample wide share of the population with sixteen years of schooling, this isolation index increases only from 0.008 to 0.016. This increase is real but modest and the average college graduate lives in a metropolitan area that is only slightly better educated than the average person without a college degree.

3 Sorting by skills: Theory

In this section, we present a simple model of urban agglomeration which provides an explanation for the tendency of skilled people to live among other skilled people. The core assumption in the model is that the number of employers depends on the number of people, and particularly, the number of skilled people in the area.² We think of this as reflecting a process where entrepreneurs come out of a city's labor force and then tend to stay generally in that area. Duranton and Puga (2001) present a more thorough discussion of location decisions of innovators in urban areas.

While we will generally be treating firms as fixed, we assume that workers are mobile and their utility in the city must be equal to their reservation utility

² The link between population size and innovations is emphasized by Kremer (1993).

which equals \bar{U} for skilled workers and \underline{U} for unskilled workers. Utility for high skill workers in city “ i ” equals: $U(W_i^H - C^H(N_i^H + N_i^L), A_i^H)$; utility for low skill workers equals $U(W_i^L - C^L(N_i^H + N_i^L), A_i^L)$. The variables W_i^L and W_i^H refer to the endogenously determined wages received in city i by low and high skill workers respectively. The variables A_i^L and A_i^H refer to the exogenous amenity levels in the city which can disproportionately attract high and low skill individuals.

The terms $C^H(N_i^H + N_i^L)$ and $C^L(N_i^H + N_i^L)$ refer to the cost of living in city i (primarily housing costs) which is a function of the sum of the total number of high skill workers (N_i^H) and the total number of low skill workers (N_i^L), and we denote this N . We allow an increase in population to have different impacts on the cost of living for both high and low skilled workers. The equations $U(W_i^H - C^H(N_i^H + N_i^L), A_i^H) = \bar{U}$ and $U(W_i^L - C^L(N_i^H + N_i^L), A_i^L) = \underline{U}$ then determine the labor supply into this city. We can use these two equalities to define implicit functions $W_i^H(A_i^H, C^H(N_i^H + N_i^L))$ and $W_i^L(A_i^L, C^L(N_i^H + N_i^L))$ which reflect the wages that firms need to pay workers to induce them to stay in a town of size N .

There are two types of firms each of which employs only one type of worker. Each firm sells its product at a price of 1 and has access to a production technology: $\theta_j^i f(Q_j)$ where j equals H or L , Q_j refers to the number of workers hired by the firm, and $f(\cdot)$ is monotonically increasing and concave. Each firm is assumed to be started by an entrepreneur who receives the excess rents from the firm. The limiting factor to entrepreneurship is assumed to be new ideas and these ideas occur stochastically in the population. The number of entrepreneurs whose firms employ high type workers is assumed to equal $h_i + \phi h_i N_i^H$. The term h_i is an exogenous constant and the term $\phi h_i N_i^H$ assumes that the amount of entrepreneurship of this type is proportional to the number of high skilled people living in this community. We are assuming that high skill workers produce a new idea with probability h_i , and with probability ϕ this new idea involves the employment of high skill workers. We also implicitly assume that there is some natural comparative advantage to this location that ensures that these firms remain in the area, or that new ideas are location-specific so that they cannot be taken elsewhere.

The number of entrepreneurs who open low skill firms equals $l_i + l_i N_i^L + (1 - \phi) h_i N_i^H$. The term l_i is again an exogenous constant. The term $l_i N_i^L$ describes the number of low skill firms started by low skill entrepreneurs. We assume that unskilled workers cannot innovate in a way that employs more skilled workers. Finally, the term $(1 - \phi) h_i N_i^H$ refers to the number of high skill entrepreneurs who start firms that employ less skilled workers. The parameter ϕ is important in the model and captures the extent to which high skill people come up with ideas that lead only to the employment of high skill workers.

The assumption that the number of firms is proportional to the number of workers of each type is the core of the model. This assumption embeds two different ideas. First, we are assuming that innovations just occur randomly to workers. We are assuming that these events are sufficiently rare in the life of most workers that they can be treated as being measure zero in the location decision, but when they occur they create the possibility for a new product that can produce sizable rents. Second, we are assuming that when a worker has a new idea, that worker stays in his home area. This fixed nature of employment can be justified

if entrepreneurs want to stay in their cities for consumption reasons or if the innovation makes use of unique aspects of their own urban environment.

This connection between population size and entrepreneurship is the key agglomeration effect in this model. Workers are both providers of labor supply and potential innovators who create labor demand. We have also assumed that unskilled workers only innovate in firms that employ other unskilled workers, but that when skilled workers innovate with probability ϕ they employ their own human capital type and with probability $1 - \phi$ they employ workers of another human capital type. We think it is reasonable to assume that people without education will have difficulty figuring out a new production process that uses entirely people with more education, but it is possible that skilled people (like Henry Ford) will occasionally develop firms that use less skilled workers.

With these assumptions the core equations of the model equate the marginal product of labor for both high skilled firms and low skilled firms with the wage that is demanded by workers to stay in that town, or:

$$\theta_i^H f\left(\frac{N_i^H}{h_i + \phi h_i N_i^H}\right) = W_i^H(A_i^H, C^H(N_i^H + N_i^L))$$

and

$$\theta_i^L f\left(\frac{N_i^L}{l_i + (1 - \phi)h_i N_i^H + l_i N_i^L}\right) = W_i^L(A_i^L, C^L(N_i^H + N_i^L)).$$

These equations do not include any classic human capital externalities (as in Lucas, 1988), and the only benefit for less skilled workers from living around the more skilled is actually a pecuniary externality. If $h_i = 0$, the model suggests that each firm has decreasing returns to skilled workers, but that the city as a whole has constant returns to skilled workers. Comparative statics then follow:

Proposition 1. *The number of skilled people in the city will decline and the number of unskilled people in the city will rise if θ_i^L , l_i or A_i^L increase.*

If $(1 - \phi)/C^L(N)$ is sufficiently high, then an increase in θ_i^H , l_i or A_i^H will cause both the number of skilled people and the number of unskilled people in the city to rise.

If $(1 - \phi)/C^L(N)$ is sufficiently low, then an increase in θ_i^H , l_i or A_i^H will cause the number of skilled people in the city to rise and the number of unskilled people in the city to fall.

Proposition 1 provides the core results of this model. Parameters that increase the attractiveness of the city for the less skilled, such as θ_i^L , l_i or A_i^L , will cause the cost of living to rise and skilled people leave the city. An increase in θ_i^L (the productivity of less skilled firms) or l_i (the number of firms specializing in less skilled workers) will create an increase in labor demand for the less skilled in that location. An increase in A_i^L reflects an amenity that particularly attracts the poor, such as more generous welfare payments, will also repel the rich as it seems to have done in cases like East St. Louis.

The more interesting results concern the comparative statics on θ_i^H , h_i and A_i^H . Because an increase in the skilled leads to more firms that hire both the skilled and the unskilled it is quite possible that a parameter that makes a place more attractive to the skilled will end up attracting more unskilled as well. The impact of the skilled on the location of the unskilled depends on the magnitude of $(1 - \phi)/C^U(N)$. The numerator of this expression captures the extent to which the skilled are likely to produce innovations that employ the unskilled. The denominator captures the extent to which an increase in the number of the skilled living in the area drives up the cost of living for the unskilled.

When $(1 - \phi)/C^U(N)$ is high, either because the skilled innovate in a way that employs the unskilled or because $C^U(N)$ is low, then we would expect the skilled and unskilled to live together. Any location specific characteristic that attracts the skilled, such as a consumption amenity (A_i^H), or the productivity of more skilled workers (θ_i^H) or an increase in the number of firms that hire the more skilled, will also attract more unskilled people as well who will choose to locate near the skilled.

However if $(1 - \phi)/C^U(N)$ is low, either because the skilled innovate in a way that only employs the skilled or because $C^U(N)$ is high, then the forces that attract the skilled will tend to push the unskilled away. In this case, the impact of the skilled on the cost of living overwhelms the impact of the skilled on the employment prospects for the less skilled. Furthermore, in this scenario we should expect to see more clustering of skilled and unskilled people in different metropolitan areas.

This proposition suggests two reasons why we should see an increasing tendency for more skilled workers to go to cities that have advantages for the skilled. First, innovations of the skilled may now tend to employ primarily skilled persons while in the past, innovative firms founded by skilled people often employed large numbers of unskilled people. In a sense this is a comparison between Henry Ford and Bill Gates. Both individuals were themselves skilled, but Ford's production involved vast numbers of unskilled workers. Gates' company primarily employs the more skilled.

A second reason why we should expect to see a change in the locational patterns of the skilled and unskilled is increasing housing supply inelasticity. When housing supply is sufficiently elastic so that $C^U(N)$ is small, then even if the innovations of the skilled only rarely employ unskilled, the unskilled will still come to the town because the cost of living there is so modest. However, when $C^U(N)$ is large then the skilled will push up housing prices for the unskilled by such a large amount that they will not be willing to pay to live around the skilled entrepreneurs. As such, a second possible reason why the skilled increasingly locate around each other is that the housing supply has gotten more inelastic over time, possibly as a result of land use regulations (Glaeser, Gyourko and Saks 2005).

We end this section by considering the impact of an increase in ϕ on wages, but in this case we hold city populations fixed. A model of the form described above has essentially perfect labor supply so that shocks to labor demand can only have a limited effect on nominal wages and no effect on real wages. This could

be modified by assuming heterogeneous tastes for different locales, but this would complicate the situation excessively. Instead, we simply assume that in some short run, city populations are fixed.

As such, this proposition should be seen as a partial equilibrium analysis that restricts population movements to changes in ϕ . If population levels are allowed to move freely with respect to changes in this parameter, then the differences in wages will be entirely determined by differences in the cost of living functions and will not reflect different demand for workers. To simplify matters further, we assume that $h_i = l_i = 0$ and find:

Proposition 2. *The difference in wages between the skilled and unskilled will increase with ϕ .*

If higher skill levels in cities are caused by differential amenity levels or higher levels of θ_i^H then as ϕ rises, the sensitivity of average income to the share of high skilled people in the area will also rise as long as f'' is effectively constant in the range of interest and $\theta_i^H > \theta_i^L h_i^2/l_i^2$.

If $f(Q) = Q^\alpha$ then the differences in the logarithm of wages between skilled and unskilled that is associated with an increase in ϕ will be greater in areas where N_i^H/N_i^L is higher.

Proposition 2 provides the basis for our analysis in Section V where we look at the changes in wages over time. The proposition has three parts. First, as ϕ rises this will mean that there are more firms using skilled labor relative to firms using unskilled labor. This causes labor demand for skilled workers to rise and this will increase the wages of the skilled relative to the unskilled. This effect is not surprising, but it relates to the well known rise in returns to skill over the last 25 years. If new firms make more use of skilled workers, then this provides one possible reason for this phenomenon.

The second part of the proposition shows that the connection between area level skills and area level income will increase as ϕ rises. The primary reason for this is that the rise in returns to skill naturally means that the relationship between area level skills and area level income will rise.

The third part of the proposition suggests that an increase in ϕ will cause greater wage gaps between skilled and unskilled in more skilled areas. The reason for this is that higher values of ϕ mean that the skilled entrepreneurs, who are more abundant in more skilled cities, generally raise the wages of skilled laborers. A higher value of ϕ essentially increases the degree of complementarity across skilled people since it means that skilled people are more likely to end up hiring other skilled people.

4 Skills, innovation and housing supply

We now turn to evidence on changes in the ratio $(1 - \phi)/C^{L'}(N)$. We start by documenting the tendency of skilled innovators to hire skilled laborers. We then turn to the changes in housing supply elasticity.

4.1 Skills and innovation

One possible explanation for skill divergence among cities is that the innovations of skilled entrepreneurs disproportionately utilize skilled workers. Although we are not aware of a publicly available data source that would allow us to test this hypothesis directly, we have assembled several pieces of evidence that are broadly consistent with such an explanation.

Perhaps the most direct evidence on the relative level of human capital in new firms comes from Abowd et al. (2001), who use a restricted access data set from the Census Bureau's Longitudinal Employer-Household Dynamics Program. Examining the Illinois economy, the authors find that one important factor contributing to the overall increase in human capital in the 1990s is that entering businesses had substantially more skilled employees than exiting businesses. They note, however, that upskilling within continuing businesses accounts for the largest share of the aggregate increase in skills. This work does not specifically relate to the difference between skilled and unskilled entrepreneurs, but if we believe that the majority of entrepreneurs are skilled, then this fact supports our assumption that ϕ has changed.

Another testable implication of the model is that the pattern of skill divergence we observe across cities also holds across industries. Using the census PUMS data, we computed average educational attainment by 3-digit industry and examined whether highly skilled industries have a tendency to become more skilled over time. Column one of Table 5 shows this relationship. For each time period, there is a positive association between the initial share of workers with a college degree and the growth in college attainment over the succeeding decade.

Table 5. Skill divergence and employment growth across industries

	1970s	1980s	1990s
Dependent Variable: Change in Percent BAs in Industry			
Initial Share of BAs in Industry	0.111 (0.026)	0.037 (0.021)	0.063 (0.025)
Observations	142	142	133
R^2	0.119	0.022	0.048
Constant	0.047	0.024	0.013
Dependent Variable: Log Change in Industry Employment			
Initial Share of BAs in Industry	1.136 (0.210)	1.48 (0.222)	0.470 (.224)
Observations	142	142	133
R^2	0.145	0.245	0.026
Constant	0.049	-0.094	-0.081

Notes: Standard errors in parentheses. Data are from the 1970, 1980, 1990, and 2000 PUMS. Educational attainment and total employment are computed for a consistent set of three-digit industry codes, using the PUMS variable *ind1950* (Ruggles et al. 2004). Each observation is one three-digit industry.

The association is only significant at the 10 percent level in the 1980s, however. The relationship between initial attainment and subsequent upskilling is largest in the 1970s, suggesting that the rate of skill divergence may have decreased over time. However, we are not able to reject the hypothesis that skill divergence in the 1990s is equal to that observed in the 1970s.

We next investigate whether employment in more skilled industries is growing over time. This fact might be understood as testing the hypothesis that skilled people are particularly innovating in recent time periods. In every decade, we see a positive association between initial college attainment in an industry and subsequent employment growth in the industry. This result is simply a re-affirmation of the common finding in the wage structure literature that between-industry demand shifts have favored college workers (e.g., Katz and Murphy 1992). The relationship is strongest in the 1980s, although we cannot reject that the coefficients for the 1980s and 1970s are equal. We can, however, reject that the relationship in 1990s is equal to either of the earlier decades, indicating that the relationship between initial attainment and subsequent employment growth has declined over time, consistent with Autor, Katz and Krueger (1998). In sum, two patterns we observed across cities are also evident across industries; namely, the tendency for highly skilled cities and industries to become more skilled over time and to add more jobs.

While the above evidence on skill and employment growth across industries is consistent with the basic hypothesis that skilled entrepreneurs disproportionately hire skilled workers, we have still not shown an explicit link between the skills of entrepreneurs and workers. Although we lack data on entrepreneurs specifically, we approach this problem by computing educational attainment by occupation within industries. Using the census PUMS, we compute the percentage of college graduates in three broad occupational categories: managerial and proprietary, technical and professional, and all other workers.³

We compute these occupational attainment estimates by 3-digit industry code, excluding farming and agriculture from the analysis. As shown in Table 6, we then correlate the educational attainment among managers and owners with the educational levels for the other groups. In 1970, across industries the correlation between manager's education levels and the education level of ordinary workers was 38 percent. In 2000, the correlation between manager's education level and the education level of ordinary workers was 51 percent. This difference is strongly significant statistically.

Figure 3 shows the relationship between the share of managers with college degrees and the share of ordinary workers with college degrees in 1970. The regression line in the graph shows that on average a one percent increase in the share of managers with college degrees is associated with a 0.15 percent increase

³ Specifically, using the PUMS occupation codes (Ruggles et al. 2004), we classify individuals into three categories: Managers, Officials, and Proprietors (codes 200 to 290); Professional and Technical (codes 000 to 099); and all other workers. We exclude farming occupations (codes 100, 123, and 810 to 840) and those with missing values from the analysis.

Table 6. Changing correlation between education of managers and workers

	Worker BA Attainment	Tech./Prof. BA Attainment	Manager/Proprietor BA Attainment
1970			
Worker BA Attainment	1.000		
Tech./Prof. BA Attainment	0.351	1.000	
Manager/Proprietor BA Attainment	0.385	0.425	1.000
2000			
Worker BA Attainment	1.000		
Tech./Prof. BA Attainment	0.504	1.000	
Manager/Proprietor BA Attainment	0.515	0.637	1.0000

Notes: Data are from the 1970 and 2000 PUMS (Ruggles et al. 2004). Workers are classified into one of the three occupational categories as described in the text, using the PUMS variable *occ1950*. Educational attainment is computed by occupational category within three-digit industry codes, using the PUMS variable *ind1950*. Each observation is a three-digit industry \times occupational category.



Fig. 3. Relation of manager education to worker education across industries, 1970

in the share of ordinary workers with college degrees. Figure 4 shows the same relationship for 2000. The regression line in that graph shows that on average a one percent increase in the share of managers with college degrees in 2000 is associated with a 0.38 percent increase in the share of ordinary workers with college degrees. In other words, high skill managers increasingly manage firms with high skill workers.

In Table 6, we also show the correlations between the education of managers and the education of professional and technical workers. This correlation coefficient was 0.42 in 1970 and 0.64 in 2000. Again, the change is statistically significant. Highly skilled managers and skilled technical workers have also increasingly concentrated within industries.

Our findings are not definitive and there are substantial questions about causality in all of our empirical findings, but our evidence suggests the idea that skilled entrepreneurs and managers disproportionately hire skilled labor to be a promising explanation for skill divergence across cities. The increasing correlation between managerial and worker skill, and the persistent tendency for initial skilled industries to become more skilled over time provides some evidence for this view. These results are also broadly consistent with Kremer and Maskin's (1996) finding of increasing segregation of high- and low-skill workers into separate firms.



Fig. 4. Relation of manager education to worker education across industries, 2000

Note for Figures 3 and 4: Data are from the PUMS. Each circle represents one 3-digit industry. Managers and non-technical workers are defined as described in the text.

4.2 Housing supply

A second possible explanation for the increasing tendency of skilled cities to become more skilled is that housing supply appears to have become more inelastic. While we lack compelling measures of supply elasticity, there are a variety of pieces of evidence that do point toward an increasing inability to build more housing cheaply. In this section, we summarize that evidence.

The first piece of evidence pointing toward this inelasticity is the increasing gap between housing prices and construction costs. In 1970, in almost every metropolitan area, most houses were valued at less than 1.25 times construction costs. Ninety percent of the areas in a 102 metropolitan area sample had an average price that was less than twelve percent more than construction cost. Thirty years later in the average metropolitan area, housing prices cost more than 46 percent more than construction costs (Glaeser, Gyourko and Saks, 2005). These prices must reflect increasing difficulties of building readily.

A second piece of evidence suggesting declining housing supply elasticity is the declining number of new homes that are being built, particularly in high cost areas. In a sample of homes in 102 metropolitan areas, the median metropolitan area in 1970 built 35 percent of its homes over the previous decade. In 2000, the median metropolitan area built only 14 percent of its homes over the previous decade (Glaeser, Gyourko and Saks, 2005). Detailed analysis of particular metropolitan areas shows a stunning decline in permitting between 1960 and today (Glaeser, Gyourko and Saks, 2004).

A third piece of evidence showing declining elasticity is that the connection between housing prices and new construction is falling. Glaeser, Gyourko and Saks (2004) show that in the 1960s and 1970s in New York City, new construction increases when housing prices rise. In the 1990s, this correlation disappears. Glaeser, Gyourko and Saks (2005) find across metropolitan areas that higher housing prices are becoming more negatively associated with new construction over time. This decline is consistent with an increasing inability to build in high cost regions.

But while there is evidence suggesting that housing supply has become more inelastic over time, there is little evidence suggesting that the housing price premium associated with living in a better educated city has increased. If we regress median housing values in 1980 on the share of adults with bachelor's degrees during that year, we estimate:

$$\text{Log}(\text{Value}_{1980}) = 10.22 + 3.012 \cdot \text{College}_{1980} \quad (8)$$

(0.042) (0.241)

Standard errors are in parentheses and there are 318 observations. The R^2 is 33 percent. Data are from the census and housing values are in 2000 dollars.

The same regression in 2000 yields:

$$\text{Log}(\text{Value}_{2000}) = 10.88 + 3.007 \cdot \text{College}_{2000} \quad (9)$$

(0.051) (0.207)

There are 318 observations and the R^2 is 40 percent. Data are from the census. There has been a tendency of housing prices to rise in highly educated areas, but generally the price increase has been lower than the price increase that would be suggested by rising wages in those areas (Glaeser and Saiz, 2004).

A second piece of evidence suggesting that high housing prices are not driving the increasing tendency of the skilled to move toward skilled areas is that controlling for lagged housing prices does little to change to relationship between initial skill levels and the growth in the share of the population with at least sixteen years of schooling. For example if we modify our basic regression from Table 1 to include the lagged median housing value, we estimate:

$$College_{2000} - College_{1990} = \underset{(0.02)}{-0.004} + \underset{(0.014)}{0.139} \cdot College_{1990} + \underset{(0.002)}{0.001} \cdot \text{Log}(Value_{1990}) \quad (10)$$

where $\text{Log}(Value_{1990})$ refers to the logarithm of median housing value in the metropolitan area in 1990. The R^2 in this regression is 28 percent and there are 318 observations. The coefficient on share with sixteen years of schooling or more is essentially unchanged from that reported in Table 1, regression 1. The coefficient on the lagged median housing value is essentially zero. There is no sense that controlling for lagged housing prices impacts the relationship between initial share of the adult population with sixteen years of schooling or more and later growth in that variable.

Overall, there is little evidence supporting the idea that increasing inelastic housing can explain the divergence of human capital levels. While housing supply is becoming more inelastic, there has been a substantially higher cost associated with living in educated metropolitan areas for many decades. As a result, it does not seem likely that high housing prices are the main reason for the increasing tendency of the skilled to live in more skilled areas.

5 Evidence on income patterns across metropolitan areas

Finally, we turn to the implications of skill agglomerations for wages and income patterns. First we turn to the correlations between area level skills and area level income. Second, we turn to patterns of metropolitan area income convergence.

5.1 Income and education across metropolitan areas

The model suggested that if the tendency of skilled entrepreneurs to employ skilled laborers is rising, we should expect an increasing correlation between area level skills and area level income. Indeed, the correlation between education and income is also increasing. Figure 5 shows the relationship between the percent of the population with a college degree and the logarithm of per capita income at the metropolitan area level in 2000. The income data come from REIS and educational attainment from the census. The regression line shown in the figure is:

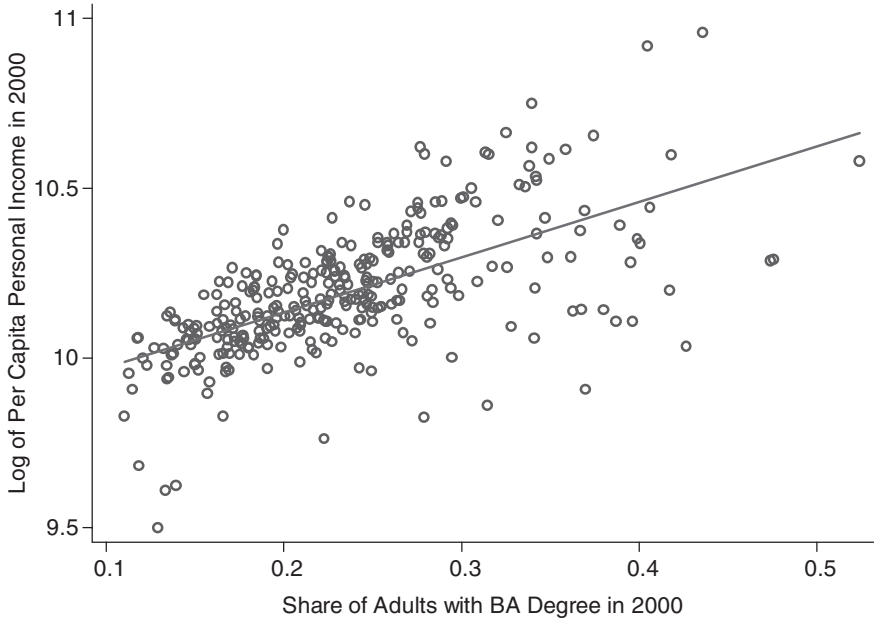


Fig. 5. Education and income, 2000

$$\text{Log}(\text{Income}_{2000}) = 9.8 + 1.63 \cdot \text{College}_{2000} \quad (11)$$

(0.02) (0.11)

The R^2 of this regression is 40 percent and there are 318 observations. In 1970, the relationship is much weaker as shown in Figure 6. The regression line in this case is:

$$\text{Log}(\text{Income}_{1970}) = 8.1 + 0.84 \cdot \text{College}_{1970} \quad (12)$$

(0.02) (0.21)

There are again 318 observations and the R^2 is 4.6 percent. The correlation between the logarithm of mean income in the metropolitan area and the percent of the population with college degrees increases from 21.4 percent in 1970 to 32.2 percent in 1980 to 49.8 percent in 1990 to 63.3 percent in 2000.

While these trends are impressive, these results in no sense correct for the general rise in returns to skill. To do this, we follow Rauch (1993) and turn to individual level regressions where we examine the changing correlation between metropolitan area level skills and individual wages holding individual skills constant. Our basic specification is:

$$\text{Log}(\text{Wage}_{i,MSA,t}) = \alpha_t \cdot \text{Schooling}_{i,t} + \beta_t \cdot \text{Schooling}_{MSA,t} + \text{Other Controls} \quad (13)$$

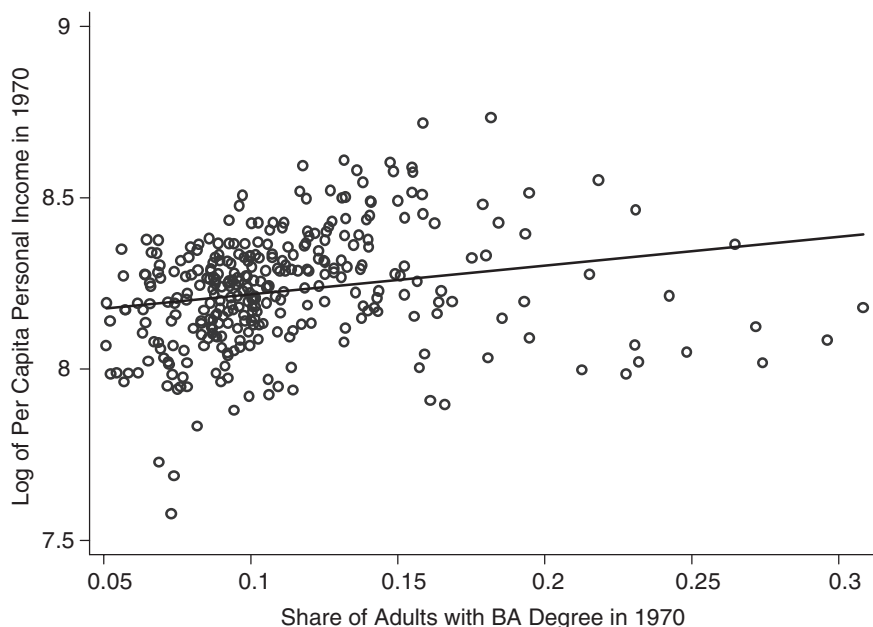


Fig. 6. Education and income, 1970

Note for Figures 5 and 6: Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England), using constant 1999 boundaries. Income data are from REIS and educational attainment from the census.

Data are from the PUMS (Ruggles et al. 2004).⁴ For each individual, the logarithm of hourly income for individual i at time t living in a particular metropolitan area ($\text{Log}(\text{Wage}_{i,MSA,t})$) is regressed on their own education and the percentage of adults with a college degree in the metropolitan area ($\text{Schooling}_{MSA,t}$). The coefficients on both individual schooling and metropolitan area level schooling are allowed to change over time. We estimate the individual schooling effects by including dummy variables for individual years of education. We include only males between 25 and 55 with complete observations. We control for race using dummy variables and for age by using a fourth order polynomial. The coefficients are also allowed to change over time.

In all of our regressions, we use average schooling in 1970 as our measure of metropolitan area schooling and estimate its effect on individual wages in 1980, 1990, and 2000. This has the effect of limiting the endogeneity of our key independent variable and of focusing on the wage effects of initially high skilled cities. If we used contemporaneous schooling levels, then the correlation with wages is almost sure to be the result of omitted variables that both raise demand for high skilled workers in the region and later attract more skilled people into that region.

⁴ For each PUMS year, we use the sample with the richest metropolitan area coverage, which is the 1 percent sample in 1970, 1980, and 1990, and the 5 percent sample in 2000 (Ruggles et al. 2004).

The results are shown in Table 7. The first equation shows our basic results, which echo those of Rauch (1993) and Glaeser and Saiz (2004). There is a strong basic connection between the share of college graduates in the metropolitan area and the logarithm of the hourly wage. Each extra percent of adults with a college degree is associated with .58 log points of extra income in 1980. This effect rises over time and in 2000, each extra percent of metropolitan area college attainment is associated with 1.22 log points of extra income. The *F*-tests reported in the bottom row of the table confirm that this change is statistically significant.

Perhaps the most significant issue with this Rauch (1993) type regression is the possibility that metropolitan area schooling is correlated with unobserved individual human capital. After all, since more skilled metropolitan areas are defined on the basis of people with more observed human capital sorting into them, it would not be surprising that people with more unobserved human capital also sort into them. If we are willing to accept that this sorting has not changed over time and that the coefficient on unobserved ability is not changing over time (both of which are quite debatable assumptions), then controlling for metropolitan area fixed effects is one way of addressing the problems related to unobserved ability sorting.

In regression (2) of Table 7, we repeat our basic specification including metropolitan area fixed effects. As such, we can no longer estimate a baseline relationship between metropolitan area years of schooling and the logarithm of hourly wages. However, we can estimate the extent to which metropolitan area schooling has become a more important predictor over time by looking at the interaction between metropolitan area schooling in 1970 and dummy variables for the years 1990 and 2000. Again, the regression shows a significant increase in the correlation between MSA level education and individual wages. The coefficient on education rises by 0.51 between 1980 and 1990 and 0.16 between 1990 and 2000. Including MSA fixed effects does little to change the basic relationships identified in regression (1).

In regressions (3)–(5), we reproduce regression (1) of Table 7 for different education subgroups (results including MSA fixed effects are quite similar). In this case, we are looking at whether there has been a change in the correlation between metropolitan area education and wages for different subgroups and this represents a test of the prediction of proposition 2 that an increase in the parameter ϕ , which represents the tendency of high human capital entrepreneurs to hire their own, will lead to higher wages particularly for high skilled workers. We divide the population into people with 16 years of education or more, between 12 and 16 years of education and strictly less than 12 years of education. This follows the work of Mare (1995) who identified different Rauch (1993) effects among education subgroups.

In regression (3), we see the changing impact of area-level education on high school dropouts. The effect only gets slightly stronger over time. The coefficient rises from 0.75 to 1.15, and we cannot reject that the coefficients for 1980 and 2000 are equal. In regression (4), we report results for high school graduates and those with some college. In this case, the basic effect is stronger and the coefficient rises between 1980 and 2000 from 0.45 to 1.08. For this group, all of the

Table 7. IPUMS estimates of effects of MSA-level education on individual wages

Total Sample	Total Sample	Did Not Graduate High School	High School Graduates, No BA	College Graduates
(1)	(2)	(3)	(4)	(5)
1970 Pct BAs × 1980 dummy	0.577 (0.262)	Omitted	0.753 (0.372)	0.454 (0.278)
1970 Pct BAs × 1990 dummy	1.131 (0.255)	0.512 (0.055)	1.181 (0.275)	0.996 (0.265)
1970 Pct BAs × 2000 dummy	1.220 (0.215)	0.669 (0.043)	1.080 (0.210)	1.414 (0.270)
MSA fixed effects	No	No	No	No
R ²	0.192	0.209	0.099	0.097
Observations	2938321	2938321	1681883	908742
F test: 2000 = 1980	0.003	0.000	0.219	0.008
F test: 2000 = 1990	0.474	0.0001	0.451	0.007
F test: 1990 = 1980	0.010	0.0000	0.109	0.120

Notes: Robust standard errors clustered by MSA in parentheses.

Data are from the 1970, 1980, 1990, and 2000 PUMS. Regressions include year fixed effects and controls for age up to a 4th degree polynomial, race, Hispanic status, and individual education. All controls are allowed to vary by year. Total sample includes all males aged 25 to 55 who worked at least 10 hours per week. Individuals with missing or imputed values for any of the variables are excluded. The dependent variable is the logarithm of the individual's hourly wage.

increase appears to have come between 1980 and 1990, as we cannot reject that the coefficients for 1990 and 2000 are equal. In regression (5), we report results for college graduates and in this case the coefficient rises from 0.68 to 1.41. The correlation between area level education and individual wages is about even for college graduates and for high school dropouts in 1970, but by 2000 a gap emerges in which the highly skilled benefit most from being around other skilled workers.

These regressions treat metropolitan area level schooling as exogenous, which may not be appropriate. After all, the major theme of the paper is the sorting of skilled people into more skilled areas. Thus, we prefer to see these empirical results as tests of one implication of the model rather than as estimates of a meaningful economic parameter.

5.2 *The end of regional convergence?*

In this final empirical section, we present what may be one of the most interesting consequences of increased sorting by skill: the decline of metropolitan area convergence. The basic convergence regression is:

$$\text{Log}\left(\frac{\text{MSA Income}_t}{\text{MSA Income}_{t-1}}\right) = \alpha + \beta \cdot \text{Log}(\text{MSA Income}_{t-1}) + \text{Other Controls} \quad (14)$$

This regression is known to be biased downwards if there is time varying measurement error in metropolitan area income. A significant literature (e.g., Barro and Sala-i-Martin, 1992) has documented a general pattern of regional convergence, and also reports a decline in the amount of convergence in the 1980s relative to earlier time periods.

In Table 8, we estimate these convergence patterns for the 1970s, 1980s and 1990s using two different measures of income: average wage and per capita income, both from REIS. In Table 8a, we include only initial income as a regressor. Both measures of income data show the same pattern. There was substantial convergence in the 1970s which decreased over time. In the case of the wage data, shown in regressions (1)–(3), the coefficient on initial income is -0.06 in the 1970s and 0.02 in the 1990s. These coefficients move from being significantly negative to being significantly positive. Figure 7 shows wage convergence in the 1970s; Figure 8 shows wage divergence in the 1990s.

Regression (4)–(6) of Table 8a shows the results for income. In this case, convergence in the 1970s is even more striking and the basic coefficient is -0.11 . In the 1990s, the coefficient estimate is positive but not statistically significant. The change in coefficients for income is even larger than the change in coefficients for wages. In the 1970s, poorer metropolitan areas were getting richer relative to richer metropolitan areas. In the 1990s, richer areas got richer relative to poorer areas.⁵

⁵ Barro and Sala-i-Martin (1999) find that convergence in personal income across U.S. states appeared to end in the 1980s, a result which they attribute to oil shocks. They do not report results for the 1990s.

Table 8a. Trends in regional convergence

	Log Change in Wages, 1970–80	Log Change in Wages, 1980–90	Log Change in Wages, 1990–2000	Log Change in Income, 1970–80	Log Change in Income, 1980–90	Log Change in Income, 1990–2000
	(1)	(2)	(3)	(4)	(5)	(6)
Initial Log Wages/Income	-0.060 (0.009)	0.025 (0.010)	0.019 (0.007)	-0.114 (0.022)	-0.062 (0.030)	0.014 (0.020)
Constant	1.798 (0.116)	0.303 (0.140)	0.254 (0.102)	1.876 (0.182)	1.195 (0.276)	0.265 (0.198)
Observations	318	318	318	318	318	318
R^2	0.13	0.02	0.02	0.08	0.01	0.00

Standard errors in parentheses.

Table 8b.

	Log Change in Wages, 1970–80	Log Change in Wages, 1980–90	Log Change in Wages, 1990–2000	Log Change in Income, 1970–80	Log Change in Income, 1980–90	Log Change in Income, 1990–2000
	(1)	(2)	(3)	(4)	(5)	(6)
Initial Log Wage/Income	-0.074 (0.008)	-0.023 (0.010)	-0.001 (0.006)	-0.150 (0.023)	-0.201 (0.022)	-0.090 (4.51)
Change in Share of Adults with College Degree	3.841 (0.591)	6.939 (0.707)	4.230 (0.463)	0.985 (0.225)	4.188 (0.226)	2.250 (10.51)
Constant	1.780 (0.109)	0.754 (0.131)	0.392 (0.092)	2.117 (0.186)	2.327 (0.201)	1.202 (6.24)
Observations	318	318	318	318	318	318
R^2	0.23	0.25	0.23	0.13	0.53	0.26

Notes: Standard errors in parentheses. Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England), using constant 1999 boundaries. Income and wage data are from REIS. Educational attainment is from the census.



Fig. 7. Regional wage convergence, 1970–1980

Note for Figures 7 and 8: Observations include all 318 metropolitan statistical areas and primary metropolitan statistical areas (NECMA definitions in New England), using constant 1999 boundaries. Data are from REIS.

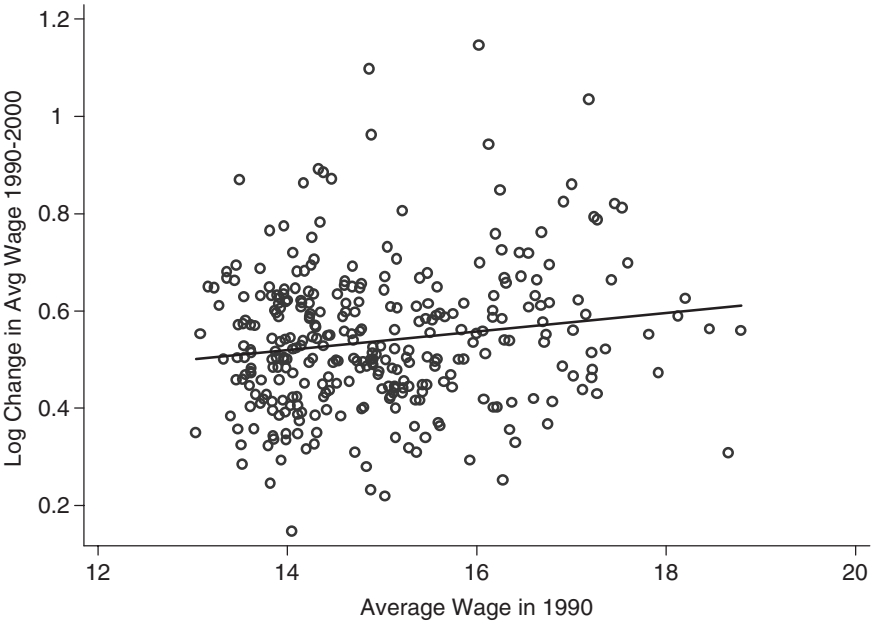


Fig. 8. Regional wage divergence, 1990–2000

One potential explanation for this fact is the rise in returns to skill and the increased sorting across metropolitan areas. In the Table 8b, we repeat our convergence regressions controlling for the change in the share of the adult population with college degrees in the metropolitan area. This control eliminates all signs of divergence. When we use log of wages in the first three regressions there is still a substantial decline in the level of convergence over the past 30 years. When we use log of income in the last three regressions, the change in the amount of convergence is more modest and it does appear that controlling for the change in the share of the population with college degrees explains the decline in regional convergence.

Convergence has declined substantially over the last 30 years. One potential reason for this is that initial high income, and high skill, places are increasingly attracting more skilled people. When we control for the changes in the skill composition at the area level, convergence reappears although there still seems to be some change between the 1970s and the 1990s.

6 Conclusion

In this article, we have documented that places with higher levels of human capital have attracted more skilled people over the last three decades. The correlation between the initial share of metropolitan area adults with college degrees and change in that variable over the 1990s is enormously strong. Educated people are still remarkably spread out across areas, but it is possible that segregation by skill might become more significant over time if this trend continues.

One potential explanation for this phenomenon is that labor demand is often created by local entrepreneurs who start firms in their own city. If skilled people are increasingly likely to start firms that hire other skilled people, then this could readily explain why an initially high level of skills would lead to a growth in the skill composition of a city over time. The initial population starts new firms and hires new skilled people. According to this view, the key change over the last 30 years is that high skill entrepreneurs increasingly innovate in ways that lead to more employment for other skilled people.

There is some evidence that such a change has occurred. The correlation between the skill level managers and workers within industries has been rising over time. The wages for skilled workers in skilled cities have been rising relative to the wages of unskilled workers in the same cities. This is a prediction of a model where the skill level in a place leads to an increase in labor demand particularly for skilled workers.

While it seems that the tendency of the skilled to move to skilled areas is partly responsible for the remarkable decline in income convergence across metropolitan areas, it is not obvious that a public policy response to this change is needed or appropriate. After all, basic results in welfare economics tell us that while wealth differences across people may be important, wealth differences across places are not necessarily troubling. Unless we are confident that the increasing tendency of skilled cities to become more skilled creates negative

externalities for the unskilled, these results suggest interesting changes across America's cities but they do not suggest any definite policy action.

References

- Abowd J, Haltiwanger J, Lane J, Sandusky K (2001) Within and between firm changes in human capital, technology, and productivity. U.S. Census Bureau, LEHD Program. Technical Paper No. TP-2001-03
- Autor D, Katz L, Krueger A (1998) Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics* 113(4): 1169–1213
- Barro R, Sala-i-Martin X (1992) Convergence. *Journal of Political Economy* 100(2): 223–251
- Barro R, Sala-i-Martin X (1999) *Economic Growth*, Cambridge, MA: MIT Press.
- Duranton G, Puga D (2001) Nursery cities: Urban diversity, process innovation and the life cycle of products. *American Economic Review* 91(5): 1454–1477
- Glaeser E, Gyourko J, Saks R (2005) Why have housing prices gone up? *American Economic Review*, forthcoming
- Glaeser E, Gyourko J, Saks R (2004) Why is Manhattan so expensive? *Journal of Law and Economics*, forthcoming
- Glaeser E, Saiz A (2004) The rise of the skilled city. *Brookings-Wharton Papers on Urban Affairs*: 47–94
- Glaeser E, Scheinkman J, Shleifer A (1995) Economic growth in a cross-section of cities. *Journal of Monetary Economics* 36: 117–143
- Katz L, Murphy K (1992) Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107(1): 35–78
- Kremer M (1993) Population Growth and technological change: One million B.C. to 1990. *Quarterly Journal of Economics* 108(3): 681–716
- Kremer M, Maskin E (1996) Wage inequality and segregation by skill. NBER Working Paper # 5718
- Lucas R (1988) On the mechanics of economic development. *Journal of Monetary Economics* 22: 3–42
- Mare D (1995) Living with graduates. Harvard University Mimeograph
- Moretti E (2004) Estimating the social returns to higher education: Evidence from cross-sectional and longitudinal data. *Journal of Econometrics* 121(1–2): 175–212
- Rauch J (1993) Productivity gains from geographic concentration of human capital: evidence from the cities. *Journal of Urban Economics* 34: 380–400
- Ruggles S, Sobek M, Alexander T, Fitch C, Goeken R, Kelly Hall P, King M, Ronnander C (2004) Integrated public use microdata series: Version 3.0 (Machine-readable database). Minneapolis, MN: Minnesota Population Center (producer and distributor)

Appendix: Proofs of Propositions

Proof of Proposition 1. Differentiating equations $U(W_i^H - C^H(N_i^H + N_i^L), A_i^H) = \bar{U}$ and $U(W_i^L - C^L(N_i^H + N_i^L), A_i^L) = \underline{U}$ gives us that:

$$\frac{\partial W_i^L}{\partial C_L} = \frac{\partial W_i^H}{\partial C_H} = 1, \quad \frac{\partial W_i^L}{\partial A_i^L} = -\frac{U_2}{U_1} < 0, \quad \frac{\partial W_i^H}{\partial A_i^H} = -\frac{U_2}{U_1} < 0.$$

For any exogenous parameter “x” that enters directly into the equation:

$$\theta_i^L f' \left(\frac{N_i^L}{l_i + (1 - \phi)h_1 N_i^H + l_1 N_i^L} \right) = W_i^L(A_i^L, C^L(N_i^H + N_i^L)),$$

but not into the equation:

$$\theta_i^H f' \left(\frac{N_i^H}{l_i + \phi h_1 N_i^H} \right) = W_i^H(A_i^H, C^H(N_i^H + N_i^L)),$$

it follows that:

$$\frac{\partial N_i^H}{\partial x} = \frac{-C^{H'}(N)}{\frac{-\theta_i^H f''(Q_H) h_i}{(h_i + \phi h_1 N_i^H)^2} + C^{H'}(N)} \frac{\partial N_i^L}{\partial x},$$

and since:

$$0 > \frac{-C^{H'}(N)}{\frac{-\theta_i^H f''(Q_H) h_i}{(h_i + \phi h_1 N_i^H)^2} + C^{H'}(N)} > -1, \text{ then } \left| \frac{\partial N_i^L}{\partial x} \right| > \left| \frac{\partial N_i^H}{\partial x} \right| \text{ with opposite sign.}$$

For $x = \theta_i^L$, differentiation yields:

$$\begin{aligned} f' \left(\frac{N_i^L}{l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L} \right) = \\ -\theta_i^L f''(Q_L) \frac{(l_i + (1-\phi)h_1 N_i^H) \frac{\partial N_i^L}{\partial \theta_i^L} - (1-\phi)h_1 N_i^L \frac{\partial N_i^H}{\partial \theta_i^L}}{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2} \\ + C^{L'}(N) \left(\frac{\partial N_i^H}{\partial \theta_i^L} + \frac{\partial N_i^L}{\partial \theta_i^L} \right). \end{aligned}$$

Assume:

$$\frac{\partial N_i^H}{\partial \theta_i^L} \geq 0 \geq \frac{\partial N_i^L}{\partial \theta_i^L},$$

and in that case both terms on the right side of the equation are non-positive and there is a contradiction so:

$$\frac{\partial N_i^L}{\partial \theta_i^L} > 0 > \frac{\partial N_i^H}{\partial \theta_i^L}.$$

For $x = l_i$,

$$\begin{aligned} -\theta_i^L f''(Q_L) \frac{N_i^L}{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2} = \\ C^{L'}(N) \left(\frac{\partial N_i^H}{\partial l_i} + \frac{\partial N_i^L}{\partial l_i} \right) - \theta_i^L f''(Q_L) \frac{(l_i + (1-\phi)h_1 N_i^H) \frac{\partial N_i^L}{\partial l_i} - (1-\phi)h_1 N_i^L \frac{\partial N_i^H}{\partial l_i}}{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2}. \end{aligned}$$

If we assume that:

$$\frac{\partial N_i^H}{\partial l_i} \geq 0 \geq \frac{\partial N_i^L}{\partial l_i},$$

the right hand side of the equation is non-positive and there is a contradiction so this proves that:

$$\frac{\partial N_i^L}{\partial l_i} > 0 > \frac{\partial N_i^H}{\partial l_i}.$$

For $x = A_i^L$,

$$\begin{aligned} -\frac{\partial W_i^L}{\partial A_i^L} = & -\theta_i^L f''(Q_L) \frac{(l_i + (1-\phi)h_1 N_i^H) \frac{\partial N_i^L}{\partial A_i^L} - (1-\phi)h_1 N_i^L \frac{\partial N_i^H}{\partial A_i^L}}{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2} \\ & + C^{L'}(N) \left(\frac{\partial N_i^H}{\partial A_i^L} + \frac{\partial N_i^L}{\partial A_i^L} \right) \end{aligned}$$

and again if we assume that:

$$\frac{\partial N_i^H}{\partial A_i^L} \geq 0 \geq \frac{\partial N_i^L}{\partial A_i^L},$$

the right hand side of the equation is non-positive and there is a contradiction so:

$$\frac{\partial N_i^L}{\partial A_i^L} > 0 > \frac{\partial N_i^H}{\partial A_i^L}.$$

For any exogenous parameter “x”, that enters directly into the equation

$$\theta_i^H f' \left(\frac{N_i^H}{h_i + \phi h_1 N_i^H} \right) = W_i^H(A_i^H, C^H(N_i^H + N_i^L)),$$

but not into the equation:

$$\theta_i^L f' \left(\frac{N_i^L}{l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L} \right) = W_i^L(A_i^L, C^L(N_i^H + N_i^L)),$$

it follows that:

$$\frac{\partial N_i^L}{\partial x} = \frac{-\theta_i^L f''(Q_L) \frac{(1-\phi)h_1 N_i^L}{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2} - C^{L'}(N)}{-\theta_i^L f''(Q_L) \frac{l_i + (1-\phi)h_1 N_i^H}{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2} + C^{L'}(N)} \frac{\partial N_i^H}{\partial x}.$$

The denominator of this expression is always positive. The numerator is positive if and only if:

$$\frac{1-\phi}{C^{L'}(N)} > \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L},$$

which since

$$\frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L}$$

is a positive finite number will always hold if:

$\frac{1-\phi}{C^{L'}(N)}$ is sufficiently high and will never hold if:

$\frac{1-\phi}{C^{L'}(N)}$ is sufficiently low.

$$\text{If } \frac{1-\phi}{C^{L'}(N)} > \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L},$$

then the signs of $\frac{\partial N_i^L}{\partial x}$ and $\frac{\partial N_i^H}{\partial x}$ must be the same and if:

$$\frac{1-\phi}{C^{L'}(N)} < \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L},$$

then the signs of $\frac{\partial N_i^L}{\partial x}$ and $\frac{\partial N_i^H}{\partial x}$ must be opposite.

$$\text{When } \frac{1-\phi}{C^{L'}(N)} < \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L},$$

then $\left| \frac{\partial N_i^H}{\partial x} \right| > \left| \frac{\partial N_i^L}{\partial x} \right|$ because

$$0 > \frac{-\theta_i^L f''(Q_L) \frac{(1-\phi)h_1N_i^L}{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2} - C^{L'}(N)}{-\theta_i^L f''(Q_L) \frac{l_i + (1-\phi)h_1N_i^H}{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2} + C^{L'}(N)} > -1.$$

When $x = \theta_i^H$, then differentiation of the first equality yields,

$$f'(Q_H) = \frac{\partial N_i^H}{\partial \theta_i^H} \left(-\theta_i^H f''(Q_H) \frac{h_i}{(h_i + \phi h_1 N_i^H)^2} + C^{H'}(N) \right) + \frac{\partial N_i^L}{\partial \theta_i^H} C^{H'}(N)$$

and if $\frac{1-\phi}{C^{L'}(N)} > \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L}$,

so $\frac{\partial N_i^L}{\partial x}$ and $\frac{\partial N_i^H}{\partial x}$ have the same sign,

then $\frac{\partial N_i^H}{\partial \theta_i^H} > 0$ and $\frac{\partial N_i^L}{\partial \theta_i^H} > 0$.

When $\frac{1-\phi}{C^{L'}(N)} < \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L}$,

then $\left| \frac{\partial N_i^H}{\partial x} \right| > \left| \frac{\partial N_i^L}{\partial x} \right|$,

so it would be impossible for $\frac{\partial N_i^L}{\partial \theta_i^H} \geq 0 \geq \frac{\partial N_i^H}{\partial \theta_i^H}$,

since that would imply that $0 > C'(N) \left(\frac{\partial N_i^H}{\partial \theta_i^H} + \frac{\partial N_i^L}{\partial \theta_i^H} \right)$ and therefore $\frac{\partial N_i^H}{\partial \theta_i^H} > 0 > \frac{\partial N_i^L}{\partial \theta_i^H}$.

When $x = h_i$, differentiation yields

$$\begin{aligned} -\theta_i^H f''(Q_H) \frac{N_i^H}{(h_i + \phi h_1 N_i^H)^2} &= \frac{\partial N_i^H}{\partial h_i} \left(-\theta_i^H f''(Q_H) \frac{h_i}{(h_i + \phi h_1 N_i^H)^2} + C^{H'}(N) \right) \\ &\quad + \frac{\partial N_i^L}{\partial h_j} C^{H'}(N) \end{aligned}$$

and if $\frac{1-\phi}{C^{L'}(N)} > \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L}$, then $\frac{\partial N_i^H}{\partial l_i} > 0$ and $\frac{\partial N_i^L}{\partial l_i} > 0$.

When $\frac{1-\phi}{C^{L'}(N)} < \frac{(l_i + (1-\phi)h_1N_i^H + l_1N_i^L)^2}{-\theta_i^L f''(Q_L)h_1N_i^L}$, then $\left| \frac{\partial N_i^H}{\partial x} \right| > \left| \frac{\partial N_i^L}{\partial x} \right|$,

so it would be impossible for $\frac{\partial N_i^L}{\partial h_i} \geq 0 \geq \frac{\partial N_i^H}{\partial h_i}$ and therefore $\frac{\partial N_i^H}{\partial h_i} > 0 > \frac{\partial N_i^L}{\partial h_i}$.

When $x = A_i^H$, differentiation yields:

$$-\frac{\partial W_i^H}{\partial A_i^H} = \frac{\partial N_i^H}{\partial A_i^H} \left(-\theta_i^H f''(Q_H) \frac{h_i}{(h_i + \phi h_1 N_i^H)^2} + C^{H'}(N) \right) + \frac{\partial N_i^L}{\partial A_i^H} C^{H'}(N),$$

$$\text{and if } \frac{1-\phi}{C^{L'}(N)} > \frac{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2}{-\theta_i^L f''(Q_L) h_1 N_i^L},$$

$$\text{then } \frac{\partial N_i^H}{\partial A_i^H} > 0 \text{ and } \frac{\partial N_i^L}{\partial A_i^H} > 0.$$

$$\text{When } \frac{1-\phi}{C^{L'}(N)} > \frac{(l_i + (1-\phi)h_1 N_i^H + l_1 N_i^L)^2}{-\theta_i^L f''(Q_L) h_1 N_i^L},$$

$$\text{then } \left| \frac{\partial N_i^H}{\partial x} \right| > \left| \frac{\partial N_i^L}{\partial x} \right|,$$

so it would be impossible for $\frac{\partial N_i^L}{\partial A_i^H} \geq 0 \geq \frac{\partial N_i^H}{\partial A_i^H}$ and therefore $\frac{\partial N_i^H}{\partial A_i^H} > 0 > \frac{\partial N_i^L}{\partial A_i^H}$.

Proof of Proposition 2. The wages of the skilled equal $\theta_i^H f'\left(\frac{1}{\phi h_1}\right)$ and the derivative of this with respect to ϕ equals $-\frac{\theta_i^H}{\phi^2 h_1} f''\left(\frac{1}{\phi h_1}\right)$ which is positive. The

$$\text{wages of the unskilled equal: } \theta_i^L f'\left(\frac{1}{(1-\phi)h_1 \frac{N_i^H}{N_i^L} + l_1}\right)$$

and the derivative of this with respect to ϕ equals:

$$\frac{\theta_i^L h_1 \frac{N_i^H}{N_i^L}}{\left((1-\phi)h_1 \frac{N_i^H}{N_i^L} + l_1\right)^2} f''\left(\frac{1}{(1-\phi)h_1 \frac{N_i^H}{N_i^L} + l_1}\right) \text{ which is negative.}$$

If z denotes the share of high skilled people in the area, then average income is:

$$\bar{Y} = z\theta_i^H f'\left(\frac{1}{\phi h_1}\right) + (1-z)\theta_i^L f'\left(\frac{1}{(1-\phi)h_1 \frac{z}{1-z} + l_1}\right).$$

Allowing θ_i^H and A_i^H to vary and differentiating totally yields:

$$d\bar{Y} = \left(\frac{\theta_i^H f' \left(\frac{1}{\phi h_1} \right) - \theta_i^L f' \left(\frac{1}{(1-\phi)h_1 z + l_1} \right)}{\left(\frac{\theta_i^L h_1 (1-\phi)}{(1-\phi)h_1 \frac{z}{1-z} + l_1} \right)^2 f'' \left(\frac{1}{(1-\phi)h_1 z + l_1} \right)} \right) dz + z f' \left(\frac{1}{\phi h_1} \right) d\theta_i^H$$

where $dz = \frac{\partial z}{\partial \theta_i^H} d\theta_i^H + \frac{\partial z}{\partial A_i^H} dA_i^H$.

The coefficient on dz represents the sensitivity of \bar{Y} to changes in z , while the coefficient on $d\theta_i^H$ represents the additional sensitivity of income for changes in z due to an increase in θ_i^H rather than A_i^H . Both coefficients will be positive provide that $\theta_i^H > \theta_i^L$ and the number of firms employing low skill workers exceeds the number of firms employing high skill workers. Differentiating the coefficient on dz with respect ϕ but holding population constant yields:

$$\alpha \left(\frac{\theta_i^H}{\phi^2 h_1} + \frac{\theta_i^L h_1 \left(h_1 (1-\phi) \frac{z(1+z)}{(1-z)^2} - l_1 \right)}{\left((1-\phi)h_1 \frac{z}{1-z} + l_1 \right)^3} \right)$$

where terms involving $f'''(.)$ have been dropped and $f''(.)$ has been assumed to be constant at $-\alpha$. Our goal is to show that this expression is positive when $\theta_i^H > \theta_i^L h_1^2 / l_1^2$ by showing that:

$$\left| \frac{\theta_i^H}{\phi^2 h_1} \right| > \left| \frac{-\theta_i^L h_1 l_1}{\left((1-\phi)h_1 \frac{z}{1-z} + l_1 \right)^3} \right|$$

The left-hand side of this inequality is decreasing in ϕ and the right-hand side is increasing in ϕ so the inequality is strictest when $\phi = 1$, giving $\theta_i^H > \theta_i^L h_1^2 / l_1^2$.

Thus when $\theta_i^H > \frac{\theta_i^L h_1^2}{l_1^2}$ the coefficient on dz in the total differential of $d\bar{Y}$ is

increasing in ϕ . To see that the coefficient on $d\theta_i^H$ is also increasing in ϕ , differentiate to obtain

$$\frac{-z}{\phi^2 h_1} f'' \left(\frac{1}{\phi h_1} \right)$$

which is positive.

If $f(Q) = Q^\alpha$, then the difference of the logarithm of wages between skilled and unskilled people equals:

$$\text{Log}\left(\frac{\theta_i^H}{\theta_i^L}\right) + (1 - \alpha) \left(\text{Log}(\phi h_1) - \text{Log}\left((1 - \phi)h_1 \frac{N_i^H}{N_i^L} + l_1\right) \right),$$

and the derivative of this with respect to ϕ equals:

$$\frac{1}{\phi} + \frac{h_1 \frac{N_i^H}{N_i^L}}{(1 - \phi)h_1 \frac{N_i^H}{N_i^L} + l_1}$$

and this is increasing with $\frac{N_i^H}{N_i^L}$.