Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

Trends in the Distributions of Income and Human Capital within Metropolitan Areas: 1980-2000

Christopher H. Wheeler
and
Elizabeth A. La Jeunesse

Working Paper 2006-055A
http://research.stlouisfed.org/wp/2006/2006-055.pdf

October 2006

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

Trends in the Distributions of Income and Human Capital within Metropolitan Areas: 1980-2000*

Christopher H. Wheeler†
Research Division
Federal Reserve Bank of St. Louis
P.O. Box 442
St. Louis, MO 63166

Elizabeth A. La Jeunesse
Research Division
Federal Reserve Bank of St. Louis
P.O. Box 442
St. Louis, MO 63166

October 12, 2006

Abstract

Human capital tends to have significant external effects within local markets, increasing the average income of individuals within the same metropolitan area. However, evidence on both human capital spillovers and peer effects in neighborhoods suggests that these effects may be confined to relatively small areas. Hence, the distribution of income gains from average levels of human capital should depend on how that human capital is distributed throughout a city. This paper explores this issue by documenting the extent to which college graduates are residentially segregated across more than 165000 block groups in 359 U.S. metropolitan areas over the period 1980-2000. Using three different metrics, we find that the segregation of college graduates rose between 1980 and 2000. We also find that cities which experienced larger increases in their levels of segregation also experienced larger increases in income inequality, although our results suggest that inequality and segregation likely influence each other.

JEL Classification: D31, R23, J10

Keywords: Human Capital Externalities, Income Inequality, Segregation

*The views expressed herein are those of the authors and do not represent the official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Corresponding author: christopher.h.wheeler@stls.frb.org.

1 Introduction

The idea that human capital accumulation imparts significant social benefits has long been espoused by economists. Besides generating internal benefits to individuals in terms of higher productivity and earnings, greater amounts of schooling may also have external effects, raising the productivity and earnings of a much broader group of individuals.[1] These effects, presumably, follow from some sort of a 'learning' mechanism, whereby workers become more productive by observing and interacting with high-skill individuals.

While certainly plausible from a theoretical perspective, a growing body of empirical work has found evidence consistent with human capital externalities, at least within locally defined markets such as metropolitan areas. Moretti (2004c), for instance, finds that plant-level measures of productivity rise with the average level of education held by workers located within the same metropolitan area. Rauch (1993) and Moretti (2004a) both find a significant link between the average levels of wages taken over all workers within a metropolitan area and the human capital possessed by the local population. Although there is somewhat less evidence that these spillover effects are present within larger geographic units such as states (e.g. Acemoglu and Angrist (2000)), this result may simply be an indication that human capital externalities have a limited geographic scope.

In fact, a recent paper by Rosenthal and Strange (2005) offers evidence that they do. Using worker-level observations from the U.S. Census, which identifies the public use microdata area (PUMA) in which each individual works, they find evidence that human capital externalities attenuate with distance. Specifically, the positive effect that the average level of human capital within a PUMA has on the wages of a given worker tends to decline as the distance between the two increases. Their results suggest that the social impacts of human

[1] As summarized by Moretti (2004b), human capital may also impart additional social benefits, including reductions in crime and greater civic participation.

capital may decrease by as much as two-thirds within a range of 15 miles.[2]

The literature on neighborhood effects (e.g. Benabou (1996), Durlauf (2004)), of course, has long drawn a similar conclusion, at least in the following sense. Given that social interactions tend to be more intense between individuals located near to one another than between individuals separated by large distances, people tend to be influenced by the composition of their neighborhoods. Case and Katz (1991), for example, find evidence of peer effects in a variety of outcomes including criminal activity, drug and alcohol use, schooling, and employment status. Although individuals are also undoubtedly affected by what happens beyond their neighborhoods, most studies of neighborhood-level outcomes suggest that individuals are strongly influenced by their immediate surroundings.[3]

How human capital externalities influence individuals within a local economy, therefore, may depend on how human capital is distributed throughout that economy. In particular, if highly educated workers tend to be widely dispersed (e.g. they are represented in many different residential areas within a local market), one would expect human capital externalities to be experienced by large fractions of the population. If, on the other hand, the highly educated tend to be concentrated in relatively few areas, the benefits may also be relatively concentrated. This line of reasoning suggests that the geographic distribution of human capital within a local economy may very well influence the relative earnings of different workers, and thus the level of income inequality the economy exhibits.

This paper examines the extent to which human capital is concentrated residentially across 359 metropolitan areas in the United States. We begin by documenting how the distribution of college-graduates by place-of-residence (defined by Census block groups) has

[2]Although not as detailed in terms of distance measures, further evidence that the spillover of knowledge may be geographically limited is provided by Jaffe et al. (1993) who find that patent citations tend to involve patents from the same state and metropolitan area.

[3]See Durlauf (2004) for a survey of empirical studies.

evolved between 1980 and 2000. We then explore the hypothesis that greater concentration of human capital is associated with greater income inequality.

Our findings indicate that, on average, college graduates became increasingly segregated from individuals with less education over this period. In 1980, the median college graduate lived in a block group with a college completion rate of 21 percent on average (i.e., the college completion rate was 21 percent or greater for at least half of all college graduates). By 2000, this individual lived in a block group with a college completion rate of 30 percent. For workers with less than a four-year college degree, the median only rose from 12 percent to 16.7 percent.[4] Three formal measures of segregation - an index of dissimilarity, an index of isolation, and a between-block group variance in college completion rates - generate qualitatively similar conclusions.

When we turn our attention to income inequality, we find a strong association between the degree of educational segregation across block groups and the degree of spread in the household income distribution. As each of our three measures of segregation rises, the level of inequality in the local income distribution increases. These estimated associations are robust to the inclusion of a variety of controls for population demographics and industrial composition as well as to the computation of educational segregation based on Census tracts rather than block groups.

Of course, our results also suggest that inequality and the residential segregation of college graduates are jointly determined: each appears to reinforce the other. It is, therefore, important to stress that our findings do not quantify the causal impact of segregation on inequality, but rather the correlation between the two. Nonetheless, the results are certainly consistent with the notion that changes in the distribution of income in the U.S. over the

[4] These statistics are calculated by taking weighted medians for each metropolitan area, where the weights are the number of college graduates and the number of individuals with less than a bachelor's degree. These particular figures represent mean values of the medians across the 359 metropolitan areas in the sample.

past several decades have been influenced by the manner in which households have sorted themselves residentially by education.

The remainder of the paper is organized as follows. The next section describes our data and the construction of our three measures of educational segregation. Section 3 presents the findings, including some basic trends in segregation and evidence on its association with income inequality. The final section concludes.

2 Data and Measurement

The data we employ comes from the decennial U.S. Census, as compiled by GeoLytics, covering a host of different geographic units, including counties, tracts, and block groups.[5] The primary advantage of the GeoLytics data relative to the summary files available directly from the Census Bureau is the consistency of the geographic units. A single set of definitions has been used in the creation of each Census variable at each level of aggregation for all years between 1980 and 2000. Hence, the change in some quantity (e.g. the rate of college completion) within a particular area is not driven by changes in the geographic coverage of that area.

To these data, we add information derived from the USA Counties 1998 on CD-ROM, which reports data on land area and rates of unemployment for all counties in the country, and unionization data from Hirsch et al. (2001). We use this latter data set to estimate metropolitan area-level rates of union coverage.[6]

In looking at residential patterns of educational segregation, we focus on block groups as our fundamental unit of analysis. We do so because block groups are relatively small: on

[5] Information about these data is available at www.geolytics.com.

[6] Hirsch et al. (2001) report state-level unionization rates. We estimate metropolitan area-level rates as weighted averages of their constituent state-level rates, where the weights are given by the fraction of each metro area's labor force located in each state. These data are available at www.unionstats.com.

average, block groups in our sample contained roughly 900 individuals of age 25 or older and covered approximately 0.33 square miles in the year 2000. As such, they constitute areas over which individuals can reasonably be expected to come into contact with one another. The notion of a human capital spillover within a block group, therefore, seems reasonable. Throughout the 359 metropolitan areas we study, there are approximately 166000 block groups (or 'neighborhoods').

An alternative geographic unit commonly used in studies of neighborhood effects, the Census tract, tends to be somewhat larger. In 2000, tracts contained 2832.4 individuals of age 25 or older and encompassed 1.3 square miles of land area, on average. We believe that these larger areas may be less meaningful as areas over which individuals can be expected to interact regularly with one another. Nonetheless, in spite of our concerns, we also conducted the analysis of educational segregation at the tract-level for the sake of gauging the robustness of our findings.

Our primary intent in this paper is to study the dispersion of individuals with and without a college degree across neighborhoods. We focus on college graduates as our population of high-human capital individuals because the empirical literature on human capital externalities has found the greatest support for spillovers emanating from this group.[7] To measure segregation, we use block group data recording the share of residents over the age of 25 with a bachelor's degree or more to construct three basic indices. The first of these measures, the index of dissimilarity, quantifies the extent to which college-educated individuals reside in the same areas of a city as individuals without a college degree:

$$\text{index of dissimilarity} = \frac{1}{2}\sum_{i=1}^{N}\left|\frac{college_i}{college_{total}} - \frac{nocollege_i}{nocollege_{total}}\right| \tag{1}$$

[7] Peri (2006), for example, finds evidence of human capital externalities stemming from the presence of college graduates, but not high school graduates.

In this expression, $college_i$ is the number of college graduates in block group i, $college_{total}$ is the total number of college graduates in the metropolitan area, $nocollege_i$ is the number of individuals without a college degree in block group i, and $nocollege_{total}$ is the total number of individuals without a college degree in the metropolitan area. This index ranges from 0 to 1 and has the interpretation as the fraction of either college graduates or non-college graduates that would have to relocate in order for the distribution of human capital to be uniform across neighborhoods (see Cutler et al. (1999)).

Because this measure may not adequately capture the extent to which people with different levels of education routinely interact with one another, we consider a second measure of segregation, the index of isolation:

$$\text{index of isolation} = \frac{\sum_{i=1}^{N}\left(\frac{college_i}{college_{total}} * \frac{college_i}{persons_i}\right) - \left(\frac{college_{total}}{persons_{total}}\right)}{\min\left(\frac{college_{total}}{persons_{min}}, 1\right) - \left(\frac{college_{total}}{persons_{total}}\right)} \tag{2}$$

As explained by Cutler et al. (1999), the term $\sum_{i=1}^{N}\left(\frac{college_i}{college_{total}} * \frac{college_i}{persons_i}\right)$ in the numerator represents the percentage of college-educated individuals in the average college graduate's neighborhood. This quantity reflects the average exposure a college graduate has to other college graduates (i.e. the average isolation from non-college graduates). We then subtract the share of college-educated people in the city as a whole to eliminate the effect of the overall size of the college-educated population. A larger percentage of the population that is college educated should increase their exposure to (or reduce their isolation from) non-college graduates. The denominator, $\min\left(\frac{college_{total}}{persons_{min}}, 1\right) - \left(\frac{college_{total}}{persons_{total}}\right)$, where $persons_{min}$ is the smallest population (25 years of age or older) among all of the city's block groups, represents the maximum possible value of the numerator.[8] Dividing by this quantity normalizes the index to values between 0 and 1.

[8]In our sample, there were sufficiently many college graduates in each metropolitan area in each year to set the denominator equal to $1 - \left(\frac{college_{total}}{persons_{total}}\right)$.

Finally, we calculate the weighted variance of college attainment among block groups:

$$\text{weighted variance} = \sum_{i=1}^{N} \left(\frac{college_i}{persons_i} - \frac{college_{total}}{persons_{total}} \right)^2 * \left(\frac{persons_i}{persons_{total}} \right) \tag{3}$$

which measures the extent of variation in college-completion shares across the neighborhoods of a given metropolitan area. We weight the variance calculation by each block group's share of total metropolitan area population (of age 25 or older) so that each block group's contribution to the variance is proportional to its relative size. Doing so mitigates the impact of extremely small block groups, which may possess extremely high or low rates of college completion, on the calculations.

Although these measures can clearly be applied to segregation based on any binary indicator, they have typically been used in studies of racial segregation (e.g. Taeuber and Taeuber (1955), Massey and Denton (1988), Cutler et al. (1999)). Still, there have been applications to segregation based on income and education. Farley (1977), for example, uses the dissimilarity index to study the segregation of individuals according to 10 different educational categories across Census tracts in 29 metropolitan areas for the year 1970.

In this study, we attempt to provide a more focused look at the segregation of high-human capital individuals (i.e. college graduates) rather than the segregation of individuals belonging to a wide array of educational groups, and we do so for a recent two-decade period rather than at a single point in time. Moreover, we seek to connect the extent of educational sorting across neighborhoods to income inequality. As far as we are aware, few existing studies have attempted to do so.

3 Empirical Results

3.1 Trends in the Segregation of College-Educated Individuals

Between 1980 and 2000, the residential distribution of college graduates became increasingly concentrated across the nation's urban areas. In 1980, the 90th percentile of the distribution of block group college completion rates averaged 30.1 percent over the 359 metropolitan areas in the sample. By 2000, this figure had reached 42.9 percent. Over this same time period, the average 10th percentile only rose from 5.8 percent to 7.1 percent.[9] Hence, even though the average college completion rate among metropolitan area populations increased from 15.7 percent to 22.5 percent, this rise in human capital was unevenly distributed across neighborhoods.

Our three formal measures of segregation of college graduates show similar qualitative trends. Summary statistics describing each measure in each of the three Census years we consider appear in Table 1. The index of dissimilarity increased by 13.3 percent, rising from 0.29 in 1980 to 0.33 in 2000. Given the standard interpretation of this measure, this result suggests that, compared to 1980, an additional 4 percent of the population with a bachelor's degree or more would have had to move in 2000 for college graduates to be uniformly distributed, on average. All of this increase, as it happens, occurred during the 1980s. Between 1990 and 2000, the index actually fell from 0.34 to 0.33.

The second measure, the index of isolation, rose by more than 50 percent over this period, increasing from 0.076 in 1980 to 0.118 in 2000. Unlike dissimilarity, isolation increased during both decades, although once again most (90 percent) of the overall increase took place during the 1980s.

The weighted variance nearly doubled, rising from 0.011 in 1980 to 0.021 in 2000. Again,

[9] These figures represent weighted percentiles where the weights are given by each block group's population 25 years of age or older.

the majority of this increase (roughly two-thirds) took place during the 1980s, although on average, the degree of dispersion in college attainment across block groups continued to rise during the 1990s.

In order to provide a better sense of the magnitudes of these measures in the context of some specific cities, Tables 2A, 2B, and 2C list the top and bottom 15 metropolitan areas according to the average values of each segregation measure between 1980 and 2000. Given that all three measures are highly correlated with one another, it is not surprising to see a fair amount of overlap among them.[10] There are, for example, seven metropolitan areas that appear among the 15 highest values of each index. State College, PA registers the highest dissimilarity (0.491) and weighted variance (0.0535), and the second-highest index of isolation (0.247). Bloomington, IN has a somewhat higher level of isolation (0.25), but ranks second in dissimilarity (0.473), and fifth according to the weighted variance (0.0473).

Among those at the bottom, five metropolitan areas appear in all three tables. Punta Gorda, FL, for example, ranks as having the lowest index of dissimilarity (0.147), the second-lowest index of isolation (0.028), and second-lowest weighted variance (0.0037). Similarly, St. George, UT has the lowest isolation (0.026), second-lowest dissimilarity (0.165), and third-lowest weighted variance (0.004).

3.2 Correlates of Segregation

What types of metropolitan areas tend to be characterized by high levels of, or large changes in, educational segregation? In this section, we attempt to sketch an answer by conducting a descriptive analysis of the residential segregation of college graduates. We do so by estimating a series of regressions of the following form:

[10] The pairwise correlations taken across all city-year observations run as follows: 0.95 between the isolation and variance series, 0.88 between the dissimilarity and isolation indices, and 0.71 between the dissimilarity and variance series.

$$y_{mt} = \mu + \delta_t + \beta X_{mt} + \epsilon_{mt} \tag{4}$$

in which y_{mt} is one of our three measures of segregation characterizing metropolitan area m in year t, μ is a constant, δ_t is a year dummy, X_{mt} is a vector of regressors, and ϵ_{mt} is a city-year-specific residual, assumed to be both heteroskedastic and correlated over time within metropolitan areas.

Because our analysis is not driven by any particular theory of educational segregation, we specify the vector of city-year-specific covariates, X_{mt}, to include a number of general characteristics: two measures of metropolitan area scale, log population and log density, three Census region dummies (West, South, and Midwest), and a number of demographic characteristics (the fraction of a metro area's population that is black, under the age of 24, over the age of 65, and the fraction of the adult population with a bachelor's degree or more). We also include nine industry employment shares (i.e. fractions of a metropolitan area's workers employed in agriculture, forestry, fisheries, and mining; construction; wholesale trade; retail trade; finance, insurance, and real estate; public administration; education services; health services; and manufacturing) to capture any association with the composition of the local economy.[11]

Results appear in Table 3A. For the most part, the estimated associations are similar (at least in terms of signs and statistical significance) across all three measures. Metropolitan areas with larger populations, for example, tend to have higher levels of educational segregation, as do cities with larger fractions of residents between 24 and 65 years of age and larger shares of blacks. This latter result may indicate that cities with larger numbers of

[11] The system by which industries are classified changed between 1990 and 2000, making the creation of a set of time-consistent industry shares difficult. We selected these nine sectors because we were able to identify them consistently across all three years.

black residents are also somewhat more segregated by race. Given that blacks tend to have lower levels of educational attainment than whites, this may produce greater segregation by college completion.[12]

The overall fraction of college graduates in the local population, interestingly, shows some variation in terms of its association with the three measures. Larger shares of college graduates in the population tend to be accompanied by greater isolation and a higher weighted variance, but less dissimilarity. Based on the preponderance of college towns in Tables 2A-2C (e.g., State College, PA; Bloomington, IN; College Station-Bryan, TX; Durham, NC; Ann Arbor, MI), one might expect to see a positive association between the overall rate of college completion and each measure of segregation. The fact that the coefficient from the dissimilarity regression is significantly negative, therefore, is somewhat surprising, although it may stem from the inclusion of the education services share in total employment.[13]

Of course, the strong, positive association between segregation and the share of education services in total employment is precisely what one would expect to see in light of the cities appearing in Tables 2A-2C. We believe that this result supports the idea that the residential segregation of college graduates is driven, to a significant degree, by the overall level of human capital within a local labor market. As college graduates become a larger and larger fraction of the population, it becomes increasingly likely that they will strongly dominate certain neighborhoods, leading to higher values of each segregation measure.[14]

[12] Because our evidence points to a rise in the segregation of college graduates between 1980 and 1990 – a decade in which the residential segregation of blacks declined (see Cutler et al. (1999)) – we do not believe that educational segregation is driven primarily by racial segregation.

[13] The two, of course, are strongly correlated. Dropping the education services share from the dissimilarity regression eliminates the significantly negative coefficient on the college share.

[14] This relationship, of course, can only hold as long as college completion rates are relatively low. That is, as the rate increases from, say, 90 percent to 100 percent, there cannot be an increase in segregation.

The estimated associations with some of the other industry shares seem to imply a similar conclusion. Greater shares of employment in construction and retail trade - industries that tend to employ relatively few college graduates - tend to be accompanied by less segregation.[15] Given that the fractions of the population either under 24 or over 65 years of age correlate negatively with rates of college completion, the negative coefficients on the two age distribution variables are also consistent with this notion.

To gain a better sense of what types of characteristics are associated with the recent *rise* in educational segregation, we also estimated (4) in 10-year differences so that the change in each measure of segregation is regressed on the changes in the vector of covariates.[16] Those results appear in Table 3B.

Although some of the variables that produced significant coefficients in the levels specification do not produce significant coefficients here (e.g. population, the share of blacks in the population, the two age distribution variables), there are some notable similarities between the two sets of results. Increasing shares of employment in education services, for example, tend to be accompanied by increasing dissimilarity and isolation, while increases in the overall share of college graduates in the population are associated with rising isolation and between-neighborhood variation in college completion. Again, these results are compatible with the idea that rising college attainment within a city's total population tends

However, given that the mean college completion rate in our sample of 359 metropolitan areas is 0.19 (with a maximum of 0.52) taken across all three years, we believe that this relationship accurately describes our data.

[15]According to the U.S. Census, 24.4 percent of individuals 25 years of age or older had a bachelor's degree or more in the year 2000. Based on our calculations using 5 percent samples from the Integrated Public Use Microdata Samples (www.ipums.org), 8.9 percent of construction workers and 13.6 percent of retail employees had a bachelor's degree in 2000.

[16]Doing so also allows us to account for any time-invariant metropolitan area-specific fixed effects that may influence the level of segregation. Because there still may be regional effects influencing 10-year differences, we included region dummies in these regressions.

to be accompanied by increases in the segregation of college graduates.

We also see some evidence that rising percentages of employment in wholesale trade and manufacturing correspond to less segregation over time. Neither of these quantities showed a significant association with levels of segregation, yet we believe that the significance of these shares further highlights the importance of aggregate human capital in driving the segregation of college graduates. While not as strikingly low as retail trade or construction, college completion rates in manufacturing and wholesale trade tend to be small when compared to the U.S. population as a whole.[17]

There are, however, two coefficients that seem to conflict somewhat with the idea that rising human capital is associated with greater segregation. Increasing shares of employment in finance, insurance, real estate and public administration are associated with lower segregation, not higher segregation. Both of these sectors tend to have relatively high rates of college completion (more than 35 percent according to our calculations). As such, while we conclude that the majority of our results point to increased rates of college completion among a city's population as a primary determinant of rising segregation of college graduates, the evidence is not unanimous on this point.

3.3 Educational Segregation and Income Inequality

As noted in the Introduction, the degree to which human capital externalities are experienced by the residents of a metropolitan area may very well depend on how highly educated individuals are distributed across residential areas. Greater concentration of human capital in certain neighborhoods should lead to the concentration of productivity gains among relatively few, whereas a more uniform distribution of college graduates should produce a more equal set of gains. That is, there should be a direct connection between the degree of spread

[17] Based on our calculations from IPUMS data, roughly 19 to 20 percent of workers in these two sectors had completed four years of college in 2000, while 24.4 percent of the adult U.S. population had done so.

in a metropolitan area's income distribution and the extent of educational segregation it displays.

This section explores this hypothesis by estimating the following:

$$\text{Inequality}_{mt} = \mu + \delta_t + \beta X_{mt} + \epsilon_{mt} \tag{5}$$

where Inequality_{mt} is a measure of overall income inequality characterizing metropolitan area m in year t, μ is a constant, δ_t is a year dummy, X_{mt} is a vector of regressors, and ϵ_{mt} is a city-year-specific residual, assumed to be both heteroskedastic and correlated over time within metropolitan areas.

We quantify income inequality, y_{mt}, using the variance of the log household income distribution. Although we would like to be able to examine additional measures of inequality (e.g. percentile differentials), we are unable to do so because we do not have household-level data on income. Instead, we have information about the numbers of households with incomes belonging to certain categories. This information allows us to approximate the variance of each metropolitan area's income distribution using a procedure outlined in the Appendix. The resulting metropolitan area income variances average 0.54 for 1980, 0.64 for 1990, and 0.65 for 2000. These magnitudes are roughly similar to those computed from microdata for the country as a whole.

We specify the vector of city-year-specific covariates, X_{mt}, to include a number of characteristics that previous work suggests may influence inequality, including indicators of race, gender, marital status, place of birth, education, age, and a variety of labor market features (industry shares, the unemployment rate, the rate of union coverage). In addition to these variables, we include each measure of educational segregation (individually rather than jointly) to determine whether the geographic distribution of college graduates within a metropolitan area is associated with the degree of inequality in its income distribution.

Results appear in Table 4. In the first three columns, we have listed the results from the estimation of (5) in levels. The last three columns report results from the estimation of (5) in 10-year differences.

Most of the findings are quite standard. Larger fractions of the population that are black, female, foreign-born, and more than 65 years of age correspond to significantly higher levels of inequality (although not changes in inequality) which is intuitive given that individuals belonging to these groups tend to have relatively low incomes. In addition, we see that larger fractions of individuals with a college degree are associated with higher inequality, which may reflect the influence of skill-biased technology. If, as suggested by Acemoglu (1999), a larger supply of highly educated labor is associated with greater usage of skill-biased technologies (e.g. computer equipment) at the workplace, we should see a positive association between overall levels of education and income dispersion.

Among the industry shares, several produce significant associations, including manufacturing, the decline of which has been widely cited as a potential cause for rising levels of wage dispersion in recent decades (e.g. Bernard and Jensen (2000)). Because the wages paid by manufacturers tend to be relatively high, particularly among workers with little experience and education, the decrease of manufacturing employment has been identified as a cause for rising inequality. The results in Table 4 are certainly consistent with that view.

We also find some evidence that increased inequality may be associated with declining union activity. Because union wage contracts tend to equalize wages across workers (e.g. Fortin and Lemieux (1997)), one would expect the relationship between rates of union coverage and inequality to be negative. Five of the six coefficients are negative, and all three coefficients from the differences regressions are statistically significant. In addition, there is a strong positive association between the rate of unemployment and the level of income inequality, which supports the idea that workers at the bottom end of the earnings distribution tend to suffer larger setbacks during economic downturns than those at the top of the distribution (e.g. Blank and Blinder (1987)).

Turning to our three measures of educational segregation, all three show significantly positive associations with the level of income inequality and two, the index of isolation and the weighted variance, are significantly correlated with the change in inequality. Although far from conclusive, this finding is, once again, consistent with the idea that human capital externalities are local. If individuals with high levels of education primarily influence others located within relatively small geographic areas (e.g. neighborhoods), a city's income distribution should be influenced by the geographic distribution of college graduates. In particular, a more extensive degree of educational segregation should concentrate human capital gains primarily among the most educated who already tend to have higher incomes than their less educated counterparts. Hence, after conditioning on a number of metropolitan area-level demographic and economic characteristics, including a measure of overall average human capital, we should find a direct connection between inequality and segregation. Our findings are certainly suggestive of such a connection.

3.4 Endogeneity and Instrumental Variables

It is important to note that, while our findings reveal a positive association between income inequality and the residential concentration of college graduates, they should not be interpreted as causal. That is, the magnitudes in Table 4 should not be interpreted as measures of how much an increase in educational segregation *causes* income inequality to rise. Indeed, it is possible that rising income inequality may induce college graduates to cluster more extensively. Were this the case, the estimated coefficients in Table 4 would represent upwardly biased estimates of the true causal effect of segregation on income dispersion.

Results from a simple statistical exercise provide some evidence that this may, in fact, be the case. In particular, we augment our regressions describing the correlates of the 10-year changes in each measure of human capital concentration to include initial values of income inequality (e.g. the change in segregation between 1980 and 1990 is regressed on the variance of income in 1980). Hence, in addition to all of the variables listed in Tables 3A and 3B, we include the initial level of income dispersion, a quantity which cannot be directly determined by the dependent variable. Our results reveal a significantly positive association in all three instances: the coefficients (standard errors) are 0.038 (0.019) for dissimilarity, 0.04 (0.01) for isolation, and 0.012 (0.002) for the weighted variance. These findings are robust to the inclusion of the initial share of college graduates in the city as well as the contemporaneous change in the variance of the income distribution. Hence, higher levels of income variation tend to be associated with larger subsequent increases in the segregation of college graduates.

While certainly not conclusive, this evidence suggests that, conditional on a host of other potential determinants, cities with higher levels of income inequality tend to see their college-educated populations segregate more over the next 10 years. This finding suggests that the coefficients in Table 4 are biased upward. Shocks to income inequality (the dependent variable in equation (5)) are positively associated with the segregation measures.

When we perform an analogous exercise for income inequality, however, we find that the causality between segregation and income inequality appears to run both ways. Augmenting the regressions describing the 10-year changes in inequality with initial values of each segregation measure, we find significantly positive coefficients in each case: 0.13 (standard error = 0.02) for dissimilarity, 0.17 (0.03) for isolation, and 0.73 (0.14) for the weighted variance. Thus, even after having accounted for all of the covariates in Table 5, we find that higher levels of segregation tend to be followed by larger increases in income inequality over the next decade. In spite of the likely endogeneity of our educational segregation measures in the inequality regressions, then, we believe that the evidence still suggests that the segregation of high human capital individuals has a positive effect on income inequality.

Results from instrumental variables (IV) estimation of (5) are generally consistent with this conclusion, although they are not particularly strong. As an instrument for educational segregation, we use the number of rivers and streams running through a metropolitan area. This variable is based upon the intuitive notion that certain topographical features, including bodies of water, may encourage individuals with different characteristics to segregate more than they otherwise would. This particular quantity, of course, has been used to instrument for racial segregation (Cutler and Glaeser (1997)) and, more famously, the structure of local jurisdictions within metropolitan areas (Hoxby (2000), Rothstein (2004)).[18] While not ideal, this variable is plausibly exogenous with respect to income inequality and does show some marginal association with our three segregation measures. When we regress the level of segregation on the total number of rivers running through a metropolitan area, in addition to the variables included in the estimation of (5), we find coefficients (p-values) of 0.00002 (0.25) for dissimilarity, 0.00002 (0.06) for isolation, and 0.000004 (0.05) for the weighted variance.[19]

Our results reveal positive, albeit insignificant, coefficients on segregation in our in-

[18]Given the human capital interpretation above, we also considered the presence of a land grant college or university (based on the acts of 1862 and 1890) as a instrument for educational segregation. This variable, however, showed little relevance as it produced insignificant coefficients in our first-stage regressions. We believe that this result stems from the presence of both the overall college fraction and the share of employment in education services in those regressions.

[19]We obtained these data from the appendix that accompanies Rothstein (2004). Technically, our instrument is the number of streams running through a metropolitan area rather than the number of stream mouths, although the results are similar with either variable. Because that paper uses a different set of metropolitan area definitions, we only use a total of 272 metropolitan areas for which the definitions (based on county-level compositions) were reasonably close.

equality regressions. In particular, we find a coefficient (standard error) of 0.89 (0.92) for dissimilarity, 0.83 (0.74) for isolation, and 4.9 (4.5) for the variance. Given that the magnitudes of the estimates are roughly in-line with those of our OLS estimates reported in Table 4, we believe that these results are generally supportive of our conclusion that the segregation of college graduates tends to produce greater income inequality.

We also attempted to instrument for the *change* in segregation. Unfortunately, the number of rivers in a metropolitan area is virtually uncorrelated with the change in our three segregation measures. None of the first-stage coefficients differed significantly from zero. Hence, although we still find mostly positive IV estimates in the regressions of the change in inequality on the change in segregation (0.014 (s.e. = 1.2) for dissimilarity, 0.05 (4.4) for isolation, -0.1 (8.9) for the variance), we are unable to draw any strong conclusions from them.

3.5 Within- and Between-Neighborhood Inequality

Although the results shown thus far are consistent with the idea that human capital externalities are confined to relatively small areas, they offer relatively little insight into the mechanism by which inequality and the residential segregation of college graduates are positively related. In theory, of course, a more uneven distribution of college graduates should increase income inequality by restricting human capital spillovers to the highly-educated (i.e. the residents of neighborhoods with large numbers of college graduates). By confining less-educated individuals to neighborhoods with few college graduates, individuals with relatively low incomes would experience little benefit from aggregate human capital.

In essence, then, a more uneven distribution of college graduates should be associated with greater income inequality because it generates greater *between*-neighborhood income differences rather than larger *within*-neighborhood differences. Our data allow us to test this particular implication. We do so by first noting that the variance of the log household income distribution within a metropolitan area, σ^2, can be estimated as

$$\sigma^2 = \frac{1}{H}\sum_{i=1}^{N}\sum_{h=1}^{H_i}(y_{h,i} - \bar{y})^2 \tag{6}$$

where $y_{h,i}$ is the income of household h of neighborhood i, $\bar{y}$ is the mean household income for the entire metropolitan area, H_i is the total number of households in neighborhood i, N is the total number of neighborhoods, and H is the total number of households, $\sum_i H_i$.[20] The right-hand-side of this expression can be re-written as the sum of two terms:

$$\sigma^2 = \frac{1}{H}\sum_{i=1}^{N}\sum_{i=1}^{H_i}(y_{h,i} - \bar{y}_i)^2 + \frac{1}{H}\sum_{i=1}^{N}\sum_{i=1}^{H_i}(\bar{y}_i - \bar{y})^2 \tag{7}$$

where $\bar{y}_i$ represents the mean household income in neighborhood i. The first term, of course, is the 'within-neighborhood' component, which measures the degree of income dispersion among households within the same neighborhood. The second term, the 'between-neighborhood' component, captures the amount of income variation across different neighborhoods. Because we do not have data on individual households, we are unable to calculate the within-neighborhood components directly. Instead, we use our estimates of the overall variance of the income distribution and the between-neighborhood components to back out the within-neighborhood parts.[21] Such a procedure is similar to the one employed by Jargowsky (1996) in a study of economic segregation.

[20] The average numbers of households per metropolitan area are relatively large: 180164.6 for 1980, 208780.9 for 1990, 240407.2 for 2000. Across all three years, the minimum number of households is 8681. Hence, the difference between using a factor of $\frac{1}{H}$ in (6) instead of $\frac{1}{H-1}$ is extremely small.

[21] Within-neighborhood variation averaged 0.47 in 1980, 0.5 in 1990, and 0.52 in 2000 (with a minimum of 0.37 for all three years). Between-neighborhood variation averaged 0.07 in 1980, 0.135 in 1990, and 0.13 in 2000.

After constructing these two different inequality measures, we re-estimate (5) in both levels and differences, using each measure as the dependent variable. The resulting coefficients from each of our three segregation measures appear Table 5. In the interest of brevity, we have suppressed all of the other coefficients from the regressions.

The estimates clearly indicate that the positive association between overall income inequality and the segregation of college graduates operates through the extent of income variation between block groups. Dissimilarity, isolation, and the weighted variance all produce significantly positive coefficients in both the levels and differences specifications. At the same time, there is some evidence, albeit somewhat weaker, that the degree of within-neighborhood income variation decreases as the concentration of human capital becomes more pronounced. All three coefficients from the differences specification are significantly negative, and one of the coefficients from the levels specification is.

Such results, it should be noted, could also be viewed rather simply as a reflection of income differences between the highly educated and the less educated. That is, because college graduates tend to earn more than individuals with less education, we *should* observe an increase in between-neighborhood income variation as college graduates become more segregated. Furthermore, because a rise in segregation should also lead to a decrease in the amount of educational heterogeneity within neighborhoods, we should see a decrease in the amount of within-neighborhood income variation. In other words, the fact that rising segregation tends to be associated with increasing between-neighborhood income inequality, but decreasing within-neighborhood income inequality, need not imply anything about human capital externalities. It may simply represent the re-organization of an existing set of households from a less segregated allocation to a more segregated one.

There is one aspect of the data, however, that suggests that this hypothesis is incomplete: the total variance of the household income distribution tends to increase as the extent of educational segregation becomes more pronounced. Again, although our results largely reflect correlations between these two variables, the evidence in Section 3.4 still suggests that segregation has a positive effect on inequality. If rising segregation of college graduates simply represented the re-organization of residents in a city from a less-segregated set of neighborhoods to a more-segregated set, the variance of the household income distribution ought to remain constant. Hence, the rise in the extent of between-neighborhood income variation should be offset by a decrease in the amount of within-neighborhood variation. Empirically, however, the former greatly outweighs the latter, indicating that there is some mechanism associated with the segregation of college graduates that boosts the incomes of certain households relative to others. Neighborhood-specific human capital spillovers are certainly a plausible mechanism.

3.6 Tract-Level Results

In this section, we consider an alternative geographic unit for the measurement of neighborhood segregation, Census tracts. While we prefer block groups because they are smaller and, thus, potentially more appropriate when considering geographic areas over which individuals may come into contact with one another, tracts are a common level of aggregation in studies of neighborhood-level outcomes. Thus, for the sake of comparability with previous work, as well as testing the robustness of our findings with respect to a different level of spatial aggregation, we repeat the analysis defining neighborhoods by Census tracts.

Summary statistics describing the three tract-based measures of educational segregation appear in Table 6. From them, it is apparent that average levels of segregation across Census tracts increased between 1980 and 2000. Yet, the magnitudes of these increases were not as large as those observed among block groups. The index of dissimilarity based on tracts, for example, rose from 0.289 to 0.294 between 1980 and 2000, whereas among block groups, it increased from 0.293 to 0.332. Similarly, increases in tract-based isolation and the variance of the college completion rate averaged 0.019 and 0.007 (or roughly 26 percent and 70 percent), respectively. For block groups, they increased by 0.042 (55 percent) and 0.01 (91 percent).

Given that tracts consist of block groups, the fact that block-group level segregation exceeds tract-level segregation is quite intuitive. After all, if college graduates become more segregated across tracts, they will necessarily become more segregated across block groups. The reverse, however, is not always true. It is possible for segregation to increase across block groups while tract-level populations remain unchanged. The fact that college graduates became more segregated across block groups than across tracts simply indicates that sets of block groups within the same tract became more heterogeneous between 1980 and 2000.

Do increases in tract-based measures of educational segregation correlate significantly with income inequality? Table 7 reports the coefficient estimates for each of our tract-level measures from the estimation of (5) in both levels and 10-year differences. We have, once again, suppressed all of the other coefficients in an effort to be concise. Most are very similar to what is reported in Tables 4 and 5. All three measures of segregation are positively and significantly related to both the levels of and changes in overall inequality.[22] Moreover, this relationship clearly operates through a strong positive association with the extent of between-tract income variation, despite a negative correlation with the amount of within-tract variation.

[22] Our instrumental variables estimates are virtually identical to those obtained using block groups. We find coefficients (standard errors) of 0.9 (0.92) for dissimilarity, 0.83 (0.74) for isolation, 4.9 (4.5) for the weighted variance in levels; 0.19 (1.3) for dissimilarity, 0.65 (4.7) for isolation, and -1.2 (8.7) for the weighted variance in differences. Again, these results are based upon a reduced sample of 272 metropolitan areas.

4 Concluding Discussion

This paper has documented trends in the extent of segregation by college attainment among 359 U.S. metropolitan areas from 1980 to 2000. Our findings indicate that, on average, the residential distribution of college graduates became more uneven over these two decades.

In terms of the correlates of segregation, we find evidence that cities with large levels of (and greater changes in) their overall rates of college completion tend to have greater segregation. This may indicate that, with low rates of college completion, it is difficult for college graduates to segregate themselves from those with less education because they are not sufficiently numerous to be the dominant group within any neighborhood. As their numbers within a metropolitan area rise, however, they are more able to do so.

Of course, the extent of segregation need not increase as the fraction of college graduates in a city rises. In particular, if highly educated individuals allocate themselves throughout the neighborhoods of a metropolitan area, the degree to which they are geographically concentrated may very well decrease as they become more heavily represented. The fact that the data indicate that they do become more segregated might be a reflection of Tiebout (1956) sorting. College graduates may, for example, have preferences for certain public goods (e.g. good schools) that their less educated counterparts do not value as highly. Alternatively, given their higher incomes, college graduates may bid up the cost of living in certain particularly desirable neighborhoods, forcing those with less education to reside in different parts of the city.

This hypothesis could also be consistent with the evidence reported in Section 3.4 that higher income inequality tends to be associated with rising educational segregation subsequently. If the extent of income inequality reflects the degree of variation in preferences for public goods, one would expect to see a greater propensity for the highly educated to live amongst themselves in high inequality cities.

At the same time, our evidence also suggests that the extent to which college graduates are residentially segregated influences income inequality, which is compatible with the idea that human capital externalities are particularly concentrated within small areas. One aspect of the well-documented rise in income inequality over the past several decades, therefore, may involve the geographic distribution of human capital within the nation's residential areas. Thus, as suggested by the literature on neighborhood effects, the trend in U.S. inequality may have been influenced by the residential location decisions of the country's households.

Table 1: Summary Statistics - Segregation of College Graduates

Year	Variable	Mean	Standard Deviation	Minimum	Maximum
1980	Index of Dissimilarity	0.293	0.07	0.03	0.49
	Index of Isolation	0.076	0.04	0	0.24
	Weighted Variance	0.011	0.01	0	0.04
1990	Index of Dissimilarity	0.343	0.06	0.19	0.51
	Index of Isolation	0.113	0.04	0.03	0.27
	Weighted Variance	0.018	0.01	0	0.06
2000	Index of Dissimilarity	0.332	0.06	0.19	0.47
	Index of Isolation	0.118	0.05	0.03	0.26
	Weighted Variance	0.021	0.01	0	0.06

Note: Statistics taken across 359 metropolitan areas.

Table 2A: Metropolitan Areas with Highest and Lowest

Index of Dissimilarity, College Graduates

Top 15 Metro Areas	Average Value 1980-2000	Bottom 15 Metro Areas	Average Value 1980-2000
State College, PA	0.491	Hinesville-Fort Stewart, GA	0.232
Bloomington, IN	0.473	Holland-Grand Haven, MI	0.232
College Station-Bryan, TX	0.472	Lincoln, NE	0.228
Birmingham-Hoover, AL	0.456	Bend, OR	0.226
Auburn-Opelika, AL	0.453	Elizabethtown, KY	0.22
Durham, NC	0.447	Santa Rosa-Petaluma, CA	0.218
Trenton-Ewing, NJ	0.438	Prescott, AZ	0.215
Columbus, OH	0.436	Sheboygan, WI	0.213
Houston-Sugar Land-Baytown, TX	0.434	Mount Vernon-Anacortes, WA	0.212
San Antonio, TX	0.43	Lewiston, ID-WA	0.212
Macon, GA	0.429	Wenatchee, WA	0.204
Champaign-Urbana, IL	0.428	Coeur d'Alene, ID	0.183
El Paso, TX	0.428	Barnstable Town, MA	0.172
Blacksburg-Christiansburg-Radford, VA	0.426	St. George, UT	0.165
Morgantown, WV	0.422	Punta Gorda, FL	0.147

Table 2B: Metropolitan Areas with Highest and Lowest

Index of Isolation, College Graduates

Top 15 Metro Areas	Average Value 1980-2000	Bottom 15 Metro Areas	Average Value 1980-2000
Bloomington, IN	0.25	Hot Springs, AR	0.044
State College, PA	0.247	Altoona, PA	0.044
Durham, NC	0.235	Lewiston-Auburn, ME	0.043
College Station-Bryan, TX	0.232	Barnstable Town, MA	0.043
Trenton-Ewing, NJ	0.225	Redding, CA	0.043
Champaign-Urbana, IL	0.22	Wenatchee, WA	0.042
Ann Arbor, MI	0.217	Sheboygan, WI	0.042
Lafayette, IN	0.213	Lewiston, ID-WA	0.042
Auburn-Opelika, AL	0.209	Hinesville-Fort Stewart, GA	0.041
Blacksburg-Christiansburg-Radford, VA	0.207	Mount Vernon-Anacortes, WA	0.04
Athens-Clarke County, GA	0.2	Weirton-Steubenville, WV-OH	0.038
Houston-Sugar Land-Baytown, TX	0.195	Elizabethtown, KY	0.037
Birmingham-Hoover, AL	0.194	Coeur d'Alene, ID	0.032
Columbus, OH	0.192	Punta Gorda, FL	0.028
Ithaca, NY	0.19	St. George, UT	0.026

Table 2C: Metropolitan Areas with Highest and Lowest

Weighted Variance, College Graduates

Top 15 Metro Areas	Average Value 1980-2000	Bottom 15 Metro Areas	Average Value 1980-2000
State College, PA	0.0535	Morristown, TN	0.0052
Ann Arbor, MI	0.0526	Longview, WA	0.0052
Durham, NC	0.0507	Goldsboro, NC	0.0052
College Station-Bryan, TX	0.0473	Ocala, FL	0.0051
Bloomington, IN	0.0471	Lewiston-Auburn, ME	0.0048
Champaign-Urbana, IL	0.0464	Bay City, MI	0.0048
Ithaca, NY	0.0459	Vineland-Millville-Bridgeton, NJ	0.0047
Trenton-Ewing, NJ	0.0458	Monroe, MI	0.0045
Athens-Clarke County, GA	0.0414	Coeur d'Alene, ID	0.0045
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.0411	Altoona, PA	0.0045
Lafayette, IN	0.0402	Hinesville-Fort Stewart, GA	0.0044
Gainesville, FL	0.0399	Elizabethtown, KY	0.0042
Ames, IA	0.0398	St. George, UT	0.004
Auburn-Opelika, AL	0.039	Punta Gorda, FL	0.0037
Charlottesville, VA	0.039	Weirton-Steubenville, WV-OH	0.0035

Table 3A: Segregation Results

Levels

Variable	*Dissimilarity*	*Isolation*	*Variance*
Log Population	0.02*	0.015*	0.002*
	(0.003)	(0.002)	(0.0004)
Log Pop. Density	0.001	0.0002	0.0001
	(0.003)	(0.002)	(0.0003)
% Black	0.08*	0.06*	0.01*
	(0.025)	(0.015)	(0.002)
% College	-0.16*	0.16*	0.08*
	(0.05)	(0.04)	(0.008)
% Under 24 Years	-0.22*	-0.19*	-0.04*
	(0.12)	(0.07)	(0.01)
% Over 65 Years	-0.43*	-0.23*	-0.04*
	(0.12)	(0.08)	(0.01)
% Ag.-For.-Fish.-Min.	0.1	0.06	0.01
	(0.11)	(0.06)	(0.01)
% Construction	-0.47*	-0.19*	-0.04*
	(0.16)	(0.09)	(0.02)
% Wholesale Trade	0.38*	0.19*	0.001
	(0.23)	(0.12)	(0.02)
% Retail Trade	-0.47*	-0.25*	-0.05*
	(0.14)	(0.09)	(0.01)
% FIRE	0.08	0.08	0.01
	(0.14)	(0.1)	(0.02)
% Public Admn.	-0.15	-0.08	-0.02*
	(0.1)	(0.06)	(0.01)
% Educ. Serv.	0.77*	0.55*	0.09*
	(0.14)	(0.1)	(0.02)
% Health Serv.	-0.15	-0.11	-0.02
	(0.14)	(0.09)	(0.02)
% Manufacturing	-0.08	-0.04	-0.01
	(0.07)	(0.04)	(0.01)
West Region	0.007	0.001	-0.0004
	(0.008)	(0.005)	(0.001)
South Region	0.05*	0.029*	0.004*
	(0.007)	(0.005)	(0.001)
Midwest Region	0.026*	0.017*	0.003*
	(0.007)	(0.004)	(0.001)
R^2	0.61	0.76	0.85

Note: Dependent variables are the three measures of segregation. All regressions include time-specific fixed effects. Heteroskedasticity-consistent standard errors, adjusted for correlation within metropolitan areas, are reported in parentheses. An asterisk (*) denotes significance at 10 percent or better.

Table 3B: Segregation Results

Differences

Variable	*Dissimilarity*	*Isolation*	*Variance*
Log Population	0.007	0.003	0.003
	(0.026)	(0.01)	(0.002)
Log Pop. Density	0.025	-0.005	-0.005*
	(0.02)	(0.01)	(0.002)
% Black	0.12	0.04	0.008
	(0.13)	(0.06)	(0.01)
% College	-0.2*	0.21*	0.12*
	(0.09)	(0.06)	(0.01)
% Under 24 Years	0.12	0.04	0.016
	(0.17)	(0.07)	(0.014)
% Over 65 Years	0.19	-0.07	-0.02
	(0.18)	(0.07)	(0.014)
% Ag.-For.-Fish.-Min.	-0.2	-0.11	-0.016
	(0.15)	(0.07)	(0.012)
% Construction	-0.23*	-0.16*	-0.04*
	(0.11)	(0.05)	(0.01)
% Wholesale Trade	-0.32*	-0.2*	-0.03*
	(0.14)	(0.07)	(0.01)
% Retail Trade	0.25*	0.09	0.01
	(0.1)	(0.05)	(0.01)
% FIRE	-0.56*	-0.26*	-0.03*
	(0.17)	(0.08)	(0.015)
% Public Admn.	-0.24*	-0.12*	-0.03
	(0.11)	(0.05)	(0.01)
% Educ. Serv.	0.31*	0.18*	0.016
	(0.16)	(0.09)	(0.02)
% Health Serv.	-0.05	-0.001	-0.004
	(0.15)	(0.07)	(0.015)
% Manufacturing	-0.04	-0.07*	-0.013*
	(0.08)	(0.037)	(0.007)
West Region	0.01*	0.007*	0.001*
	(0.005)	(0.002)	(0.0004)
South Region	-0.006*	0.003*	0.001*
	(0.003)	(0.002)	(0.0003)
Midwest Region	0.004	0.002	0.0004
	(0.004)	(0.002)	(0.0004)
R^2	0.58	0.61	0.55

Note: Dependent variables are the three measures of segregation. All regressions include time-specific fixed effects. Heteroskedasticity-consistent standard errors, adjusted for correlation within metropolitan areas, are reported in parentheses. An asterisk (*) denotes significance at 10 percent or better.

Table 4: Regression Results - Income Inequality

Variable	*Levels* I	*Levels* II	*Levels* III	*Differences* I	*Differences* II	*Differences* III
Log Population	-0.006	-0.008*	-0.007*	-0.09*	-0.09*	-0.09*
	(0.004)	(0.004)	(0.004)	(0.03)	(0.03)	(0.03)
Log Pop. Density	-0.0009	-0.0007	-0.0007	-0.005	-0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.02)	(0.02)	(0.02)
% Black	0.09*	0.08*	0.08*	-0.05	-0.06	-0.07
	(0.03)	(0.026)	(0.026)	(0.1)	(0.1)	(0.1)
%Female	1.04*	1.08*	1.16*	0.39	0.37	0.37
	(0.17)	(0.17)	(0.17)	(0.26)	(0.26)	(0.26)
% Married	-0.44*	-0.46*	-0.48*	-0.27*	-0.26*	-0.27*
	(0.08)	(0.09)	(0.09)	(0.1)	(0.1)	(0.1)
% College	0.39*	0.26*	0.11	0.58*	0.53*	0.4*
	(0.07)	(0.07)	(0.07)	(0.09)	(0.08)	(0.09)
% Under 24 Years	-0.11	-0.08	-0.08	0.12	0.13	0.12
	(0.08)	(0.08)	(0.08)	(0.13)	(0.13)	(0.12)
% Over 65 Years	0.24*	0.22*	0.2*	0.11	0.13	0.14
	(0.11)	(0.11)	(0.12)	(0.15)	(0.15)	(0.14)
% Foreign-Born	0.38*	0.4*	0.4*	0.24*	0.22	0.2
	(0.05)	(0.06)	(0.06)	(0.15)	(0.14)	(0.14)
% Ag.-For.-Fish.-Min.	0.17*	0.17*	0.17*	-0.02	-0.008	-0.01
	(0.07)	(0.07)	(0.08)	(0.12)	(0.11)	(0.11)
% Construction	0.09	0.07	0.08	-0.28*	-0.26*	-0.23*
	(0.13)	(0.13)	(0.13)	(0.12)	(0.12)	(0.12)
% Wholesale Trade	0.28	0.29	0.36*	-0.11	-0.08	-0.07
	(0.18)	(0.18)	(0.19)	(0.15)	(0.15)	(0.15)
% Retail Trade	-0.05	-0.05	-0.02	0.1	0.09	0.1
	(0.1)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
% FIRE	-0.11	-0.13	-0.12	-0.48*	-0.44*	-0.45*
	(0.16)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
% Public Admn.	-0.5*	-0.5*	-0.48*	-0.37*	-0.36*	-0.34*
	(0.08)	(0.08)	(0.08)	(0.13)	(0.13)	(0.13)
% Educ. Serv.	-0.26*	-0.33*	-0.31*	-0.33*	-0.36*	-0.34*
	(0.12)	(0.13)	(0.12)	(0.14)	(0.13)	(0.13)
% Health Serv.	-0.04	-0.02	-0.02	0.02	0.02	0.02
	(0.09)	(0.1)	(0.1)	(0.12)	(0.12)	(0.12)
% Manufacturing	-0.18*	-0.18*	-0.17*	-0.34*	-0.33*	-0.33*
	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.07)
% Union Coverage	0.003	-0.008	-0.02	-0.09*	-0.09*	-0.09*
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Unemployment Rate	0.68*	0.69*	0.68*	0.48*	0.48*	0.48*
	(0.1)	(0.1)	(0.1)	(0.08)	(0.08)	(0.08)
Dissimilarity	0.28*	–	–	0.022	–	–
	(0.04)			(0.034)		
Isolation	–	0.51*	–	–	0.22*	–
		(0.08)			(0.07)	
Variance	–	–	2.73*	–	–	1.55*
			(0.49)			(0.4)
R^2	0.81	0.81	0.81	0.75	0.75	0.75

Note: Dependent variable is variance of log household income. All regressions include region- and time-specific fixed effects. Heteroskedasticity-consistent standard errors, adjusted for correlation within metropolitan areas, are reported in parentheses. An asterisk (*) denotes significance at 10 percent or better.

Table 5: Within- and Between-Neighborhood Inequality Results

	Within-Neighborhood		*Between-Neighborhoood*	
Variable	Levels	Differences	Levels	Differences
Dissimilarity	-0.084*	-0.21*	0.37*	0.23*
	(0.034)	(0.04)	(0.03)	(0.04)
Isolation	-0.09	-0.36*	0.6*	0.58*
	(0.06)	(0.09)	(0.07)	(0.08)
Variance	-0.34	-1.34*	3.08*	2.89*
	(0.34)	(0.47)	(0.42)	(0.47)

Note: Dependent variables are within-neighborhood income variation and between-neighborhood income variation. Regressions include all variables listed in Table 4 as well as region- and time-specific fixed effects. Heteroskedasticity-consistent standard errors, adjusted for correlation within metropolitan areas, are reported in parentheses. An asterisk (*) denotes significance at 10 percent or better.

Table 6: Summary Statistics - Segregation of College Graduates

Tract-Level Results

Year	Variable	Mean	Standard Deviation	Minimum	Maximum
1980	Index of Dissimilarity	0.289	0.07	0.026	0.49
	Index of Isolation	0.074	0.042	0.001	0.24
	Weighted Variance	0.01	0.008	0.0001	0.045
1990	Index of Dissimilarity	0.304	0.06	0.16	0.5
	Index of Isolation	0.09	0.044	0.02	0.25
	Weighted Variance	0.014	0.01	0.002	0.056
2000	Index of Dissimilarity	0.294	0.06	0.13	0.47
	Index of Isolation	0.093	0.046	0.016	0.23
	Weighted Variance	0.017	0.01	0.002	0.05

Note: Statistics taken across 359 metropolitan areas.

Table 7: Regression Results - Income Inequality

Tract-Level Inequaltiy and Segregation

	Overall		*Within-Neighborhood*		*Between-Neighborhood*	
Variable	Levels	Differences	Levels	Differences	Levels	Differences
Dissimilarity	0.27*	0.065*	-0.05	-0.17*	0.32*	0.24*
	(0.04)	(0.036)	(0.03)	(0.05)	(0.03)	(0.03)
Isolation	0.5*	0.24*	-0.07	-0.3*	0.57*	0.54*
	(0.08)	(0.08)	(0.06)	(0.09)	(0.06)	(0.08)
Variance	2.74*	1.67*	-0.16	-0.8*	2.9*	2.49*
	(0.51)	(0.4)	(0.37)	(0.46)	(0.4)	(0.43)

Note: Dependent variables are overall variance of log household income, within-neighborhood income variation, and between-neighborhood income variation. Regressions include all variables listed in Table 4 as well as region- and time-specific fixed effects. Heteroskedasticity-consistent standard errors, adjusted for correlation within metropolitan areas, are reported in parentheses. An asterisk (*) denotes significance at 10 percent or better.

A Appendix

We estimate the variance of a metropolitan area's income distribution as follows. For each year, the number of households with incomes falling into each of N closed intervals is reported. For 1980, there are 15 categories: 0-4999, 5000-7499, 7500-9999, 10000-12499, 12500-14999, 15000-17499, 17500-19999, 20000-22499, 22500-24999, 25000-27499, 27500-29999, 30000-34999, 35000-39999, 40000-49999, 50000-74999. For 1990, there are 24: 0-4999, 5000-9999, 10000-12499, 12500-14999, 15000-17499, 17500-19999, 20000-22499, 22500-24999, 25000-27499, 27500-29999, 30000-32499, 32500-34999, 35000-37499, 37500-39999, 40000-42499, 42500-44999, 45000-47499, 47500-49999, 50000-54999, 55000-59999, 60000-74999, 75000-99999, 100000-124499, 125000-149999. For 2000, there are 15: 0-9999, 10000-14999, 15000-19999, 20000-24999, 25000-29999, 30000-34999, 35000-39999, 40000-44999, 45000-49999, 50000-59999, 60000-74999, 75000-99999, 100000-124999, 125000-149999, 150000-199999. We use these figures to compute the fraction of households with incomes less than N distinct levels, which allows N quantiles of the household income distribution to be estimated for each metro area. For example, if 14 percent of all households have income less than 25000 dollars, we estimate the 0.14 quantile by 25000. Label these quantiles X_α. We then match these N quantiles to their corresponding values from a normal (0,1) distribution. Label these quantiles U_α. Assuming a lognormal household income distribution, X_α and U_α are related as follows:

$$X_\alpha = \exp(\zeta + U_\alpha \sigma)$$

where ζ and σ are the mean and standard deviation parameters characterizing the lognormal distribution (see Johnson and Kotz (1970, p. 117)). These parameters are readily obtained by transforming this equation logarithmically and estimating by OLS. The fit of these regressions tended to be quite high in all cases. Across the 359 metro areas, the mean adjusted R^2 was approximately 0.98 for each year, and the minimum across all metro area-year observations was 0.95. With the standard deviation, σ, the variance follows simply as σ^2.

References

Acemoglu, D. (1999) "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence." *American Economic Review.* 89, 1259-1278.

Acemoglu, D. and J. Angrist. (2000) "How Large Are Human Capital Externalities? Evidence From Compulsory Schooling Laws." NBER Macroeconomics Annual, 15.

Benabou, R. (1996) "Heterogeneity, Stratification and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review.* 86, 584-609.

Bernard, A. and J.B. Jensen. (2000) "Understanding Increasing and Decreasing Wage Inequality." *The Impact of International Trade on Wages.* R. Feenstra, ed. Chicago: University of Chicago Press.

Blank, R. and A. Blinder. (1986) "Macroeconomics, Income Distribution, and Poverty." *Fighting Poverty: What Works and What Does Not.* S. Danziger and D. Weinberg, eds. Cambridge, MA: Harvard University Press.

Case, A. and L. Katz. (1991) "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." NBER Working Paper No. 3705.

Cutler, D. and E. Glaeser. (1997) "Are Ghettos Good or Bad?" *Quarterly Journal of Economics.* 112 (3), 827-872.

Cutler, D., E. Glaeser, and J. Vigdor. (1999) "The Rise and Decline of the American Ghetto." *Journal of Political Economy.* 107 (3), 455-506.

Durlauf, S. (2004) "Neighborhood Effects." in *Handbook of Regional and Urban Economics*, Volume 4. J. Vernon Henderson and Jacques-Francois Thisse eds. New York: Elsevier, 2173-2242.

Farley, R. (1977) "Residential Segregation in Urbanized Areas of the United States in 1970: An Analysis of Social Class and Racial Differences." *Demography.* 14 (4), 497-518.

Fortin, N. and T. Lemieux. (1997) "Institutional Changes and Rising Wage Inequality: Is There a Linkage?" *Journal of Economic Perspectives.* 11, 75-96.

Hirsch, B., D. Macpherson, and W. Vroman. (2001) "Estimates of Union Density by State." *Monthly Labor Review.* 124 (7), 51-55.

Hoxby, C. (2000) "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review.* 90 (5), 1209-1238.

Jaffe, A, M. Trajtenberg, and R. Henderson. (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*. 108, 577-598.

Jargowsky, P. (1996) "Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas." *American Sociological Review*. 61 (6), 984-998.

Johnson, N. and S. Kotz. (1970) *Continuous Univariate Distributions*. Boston: Houghton Mifflin.

Massey, D. and N. Denton. (1988) "The Dimensions of Racial Segregation." *Social Forces*. 67, 281-315.

Moretti, E. (2004a) "Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross Sectional Data." *Journal of Econometrics*. 121, 175-212.

Moretti, E. (2004b) "Human Capital Externalities in Cities." in *Handbook of Regional and Urban Economics*, Volume 4. J. Vernon Henderson and Jacques-Francois Thisse eds. New York: Elsevier, 2243-2291.

Moretti, E. (2004c) "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review*. 94 (3), 656-690.

Peri, G. (2006) "Schooling Externalities, Technology and Productivity: Theory and Evidence From U.S. States." NBER Working Paper No. 12440.

Rauch, J. (1993) "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics*. 34, 380-400.

Rosenthal, S. and W. Strange. (2005) "The Attenuation of Human Capital Spillovers: A Manhattan Skyline Approach." Manuscript, Syracuse University and University of Toronto.

Rothstein, J. (2004) "Does Competition Among Public Schools Benefit Students and Taxpayers?" A Comment on Hoxby (2000)." *American Economic Review*. Forthcoming.

Taeuber, K. and A. Taeuber. (1965) *Negroes in Cities: Residential Segregation and Neighborhood Change*. Chicago: Aldine.

Tiebout, C. (1956) "A Pure Theory of Local Expenditures." *Journal of Political Economy*. 64 (5), 416-424.