Industry Localization and Earnings Inequality: Evidence from U.S. Manufacturing

Authors	Christopher H. Wheeler
Working Paper Number	2004-023B
Citable Link	https://doi.org/10.20955/wp.2004.023
Suggested Citation	Wheeler, C.H.; Industry Localization and Earnings Inequality: Evidence from U.S. Manufacturing, Federal Reserve Bank of St. Louis Working Paper 2004-023. URL https://doi.org/10.20955/wp.2004.023

Published In	Papers in Regional Science

Industry Localization and Earnings Inequality: Evidence from U.S. Manufacturing

Christopher H. Wheeler*
Research Division
Federal Reserve Bank of St. Louis
411 Locust Street
St. Louis, MO 63102
christopher.h.wheeler@stls.frb.org

March 17, 2005

Abstract

While the productivity gains associated with the geographic concentration of industry (i.e. localization) are by now well-documented, little work has considered how those gains are distributed across individual workers. This paper offers evidence on the connection between total employment and the relative wage earnings of high- and low-skill workers (i.e. inequality) within two-digit manufacturing industries across the states and a collection of metropolitan areas in the U.S. between 1970 and 1990. Using measures of overall, between-education-group, and residual inequality, I find that wage dispersion falls significantly as industry employment expands.

JEL: J31, R11
Keywords: Wage Inequality, Localization, Agglomeration Economies

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

1 Introduction

Ever since Marshall (1920) first observed that producers of a particular industry may have strong incentives to locate near one another, researchers have attempted to characterize the nature and consequences of industrial localization. Their efforts, of course, have produced a massive literature establishing a host of empirical evidence on the matter (e.g. Hoover (1948), Perloff et al. (1960), Fuchs (1962), Pred (1966), Krugman (1991), and Kim (1995), among many others).

In general, much of this work can be summarized by two broad principles. First, many industries do, in fact, exhibit substantial geographic concentration. A recent study by Ellison and Glaeser (1997), for example, finds that, of the 459 four-digit (SIC) manufacturing industries in the U.S., 446 exhibit 'excess' concentration – that is, they are more geographically localized than if their location decisions had been made at random. Second, there are significant productivity gains (i.e. 'scale effects') associated with the extent of an industry's employment in a particular locality, which provides a rationale for the reasonably high degree of localization we observe. Henderson (1986), for instance, finds a strong positive association between productivity (i.e. output per unit of input) and employment across manufacturing industries at the two-digit level in a sample of U.S. and Brazilian cities.[1] Although they vary from industry to industry, his estimates imply average elasticities in excess of 0.1 – that is, holding all other inputs constant, a doubling of an industry's own employment in a city increases a constituent producer's output by more than 10 percent,

[1] Other studies of industry productivity in local markets (e.g. cities), such as Sveikauskas (1975) and Moomaw (1981), are primarily concerned with the productivity benefits of overall city size (i.e. 'urbanization economies'), not industry localization per se.

on average.

What has not been considered for the most part, however, is the extent to which these gains in productivity are shared by all economic agents. That is, although average wage levels may rise as an industry's presence in a local market increases, do workers at all points of the earnings distribution experience similar increases? Very likely, the general neglect of this topic is related to how localization effects tend to be modeled in the literature: namely, as symmetric scale effects across all workers.[2] If true, there should be no significant association between industrial localization and the gaps between the wages of high-skill and low-skill workers (i.e. inequality). On the other hand, if localization influences the earnings of workers at different points of the wage distribution in different ways, we should observe a non-negligible association between the two.

This paper explores this issue by examining the relationship between earnings inequality and the geographic concentration of manufacturing industries. Using data from the 1970, 1980, and 1990 U.S. Census on the earnings of white male manufacturing workers across the 50 states and a set of roughly 200 metropolitan areas, I find evidence of a significantly negative association between localization and inequality. Indeed, regardless of whether inequality is defined in an 'overall' manner (raw, unconditional 90-10 wage percentile gaps), a between-education-group manner (differences between the average earnings of workers with different levels of education), or in 'residual' terms (90-10 percentile gaps in the residu-

[2]Indeed, much of the theoretical work modeling localization effects goes even farther and assumes that all labor is homogeneous (e.g. Black and Henderson (1999)). This practice, as it happens, extends beyond models of localization economies. Homogeneous labor and symmetric productivity effects also emerge in the specifications of Abdel-Rahman and Fujita (1990), Carlino and Voith (1992), and Ciccone and Hall (1996), among others, who explore productivity and spatial agglomeration more broadly.

als following regressions of wages on experience and educational attainment), the results show that increases in an industry's employment within a state or metropolitan area are associated with significant decreases in wage dispersion.

Clearly, such findings suggest that models which treat localization effects as uniform across all workers are not properly specified. Whatever mechanisms underlie the increase in productivity that accompanies industrial agglomeration tend to have a larger effect on the productivity of less-skilled workers than on the productivity of more-skilled workers. Although it may seem minor, this insight is potentially quite important because it may help to identify which theories of localization economies are plausible and which ones are not. Any viable theory of localization, quite simply, should be compatible with decreasing inequality in addition to rising productivity.

Moreover, given the decrease in the extent of localization in manufacturing in recent decades, the findings may also provide some additional insight into the rise of U.S. earnings inequality over the past 30 years. While many additional factors are undoubtedly more important in explaining this trend (e.g. decreasing union activity), declining localization may have at least contributed to the widening of the wage gaps between high- and low-skill workers.

The remainder of this paper proceeds as follows. The next section provides a brief summary of the data sources used in the analysis. Section 3 then describes the results. Section 4 offers some concluding remarks.

2 Data

The data are drawn from three primary sources. First, individual-level observations on manufacturing workers are derived from four Integrated Public Use Microdata Series (IPUMS) extracts of the U.S. Census (Ruggles and Sobek et al. (2003)): the 1970 1 Percent Form 1 State Sample, the 1970 1 Percent Form 2 State Sample, the 1980 5 Percent State Sample, and the 1990 5 Percent State Sample. Second, data on employment for two-digit manufacturing industries across states and metropolitan areas comes from County Business Patterns (CBP) for the years 1970, 1980, and 1990 (U.S. Bureau of the Census (1971, 1982, 1992)). Third, additional state and metropolitan area characteristics (education, unemployment, broad industrial composition, and resident population) for these years are taken from the 1972 County and City Data Book (CCDB) and the 1998 USA Counties on CD-ROM (U.S. Bureau of the Census (1974, 1999)).

I conduct the analysis at two different geographic levels: states and metropolitan areas. While the former are somewhat large given the notion of a 'local' market, states are reasonably common as a unit of observation in studies of agglomeration (e.g. Carlino and Voith (1992), Ciccone and Hall (1996)). They also offer a more complete coverage of industries and time periods because they involve fewer data disclosure problems (see below). Unfortunately, because theories of localization apply more readily to smaller areas with economically (as opposed to politically) defined boundaries, I also consider a sample of metropolitan statistical areas (MSAs), New England County Metropolitan Areas (NECMAs), and consolidated metropolitan statistical areas (CMSAs) if an MSA or NECMA belongs to a CMSA.[3]

[3]A total of 275 such metropolitan areas (using definitions from 1995) exist (see the USA Counties data file

As with much of the literature on wage inequality (e.g. Katz and Murphy (1992), Juhn et al. (1993)), individual-level observations are limited to white males between the ages of 18 and 65 who worked at least 14 weeks in the previous year, were not in school at the time of the survey, and who earned at least 67 dollars per week (in 1982 dollars). Doing so confines the analysis to individuals with a reasonably strong attachment to the labor force. After having further eliminated all observations for which either a worker's state or metropolitan area of residence is not identified, I am left with a total of 1449536 observations in the state sample, 781175 observations in the city sample.[4]

Wages are defined as an individual's weekly wage and salary earnings.[5] Following previous research using these particular Census data (e.g. Autor et al. (1998), Acemoglu and Angrist (1999)), topcoded earnings are imputed as 1.5 times the topcode for the 1970 and 1980 samples, and as 210000 dollars for the 1990 sample.[6] All dollar figures are converted

for the definitions). Because MSA definitions change over time, aggregation of MSAs and NECMAs to the CMSA level greatly aids in the construction of markets with consistent definitions over time. In particular, in some instances, the Census data assign the residents of certain counties to different metropolitan areas within the same CMSA in different years simply as a result of such definitional changes. Aggregating to the CMSA level circumvents this problem. Note, for expositional purposes, I use the terms 'city' and 'metropolitan area' interchangeably.

[4]A list of the 20 two-digit manufacturing industries, along with the corresponding IPUMS industry codes, appears in Table A1 of the Appendix.

[5]Because weeks worked is reported in categorical form in the 1970 Census, I estimate it for each individual as the average of the 1980 and 1990 means for the individual's educational attainment level (i.e., no high school, some high school, high school, some college, college or more) and weeks-worked category.

[6]Topcodes only appear in rare instances, averaging approximately 0.6 percent of observations within either state- or city-industries. In fact, only 4 state-industry-years and 12 city-industry-years involve percentages in excess of 10 (and, thus, have their 90th percentiles influenced by the imputation). Dropping these observations from the analysis produced results that were nearly identical to those presented here.

to real terms using the Personal Consumption Expenditure Chain-Type Price Index of the National Income and Product Accounts.

Due to disclosure limitations, the CBP data files do not always report employment figures for all industries in all areas. While in the 1970 data, no further information is provided for these undisclosed industries, the 1980 and 1990 data do report total employment in categorical form corresponding to a range of employment values.[7] For these industries, I impute state- and city-industry employment by taking the midpoint of the reported range.

The fact that the 1970 data are somewhat more incomplete, however, produces an especially difficult problem when constructing the city-industry figures for this year. City-level observations must be constructed by aggregating employment figures at the county-level where disclosure limitations are especially prevalent. For this reason, I restrict the city-level analysis to 1980 and 1990.

Because there is no unique empirical measure of inequality, I consider three general types: (*i*) an 'overall' measure given by the difference between the 90th and 10th percentiles of overall (unconditional) weekly wage distribution, (*ii*) a collection of between-education-group measures quantified by differences in the average wages of workers belonging to five educational attainment groups (no high school, some high school, high school, some college, college or more) and (*iii*) a residual measure given by the 90-10 gap in the distribution of log wage residuals following a regression of log wages on a quartic in potential experience, four educational attainment dummies (some high school, high school, some college, college or more), and a marital status indicator. These regressions are performed separately by

[7] The employment categories given are 0-19, 20-99, 100-249, 250-499, 500-999, 1000-2499, 2500-4999, 5000-9999, 10000-24999, 25000-49999, 50000-99999, 100000 or more. The largest category, incidentally, does not appear for any of the state- or city-industries used in the analysis.

both year and state/city.[8]

To account for their potential influence on the inequality measures considered, I utilize two additional variables (other than those from the CCDB and USA Counties data files mentioned above) in the analysis. First, the unionization rate is taken from Hirsch et al. (2001), who report the proportion of each state's non-agricultural wage and salary workers who belong to a union in each of the three years considered here.[9] Second, the percentage of each state or city's population that is foreign-born is derived from the Census samples based on responses to the place-of-birth question. Summary statistics describing all variables in both the state and city samples appear in Tables A2 and A3 of the Appendix.

3 Results

3.1 Overall Inequality

The connection between localization and the first measure of inequality – differences in raw, unconditional wages – is based on estimation of the following statistical model:

$$\left(\log(w_{iat}^H) - \log(w_{iat}^L)\right) = \delta_i + \delta_a + \delta_t + \theta Z_{at} + \gamma \log(N_{iat}) + \epsilon_{iat} \tag{1}$$

where w_{iat}^H and w_{iat}^L denote the weekly wage earnings of 'high-skill' and 'low-skill' workers

[8] Because the 1990 Census sample is not random, I use the IPUMS person weights in all of the percentile calculations and regression analysis for this year.

[9] City-level unionization rates are constructed from state-level rates using the following procedure. If a city belongs to a single state, I assign the corresponding state-level rate. If a city includes counties from multiple states, I take a weighted average of the appropriate state-level rates where the weights are given by the share of the city's population residing in each state.

in industry i of area (state/city) a, in year t, which I measure using the 90th and 10th percentiles of the overall wage distribution within each area-industry-year.[10] Among the determinants of this high-skill/low-skill wage gap are a set of three fixed effects designed to pick up any unobserved differences across industries (δ_i), local markets (δ_a), and years (δ_t);[11] a vector of area-specific features Z_{at} (e.g. unionization rates, local human capital), which may influence between-skill-group wage differentials; and the logarithm of industry i's total employment in the area, N_{iat}, which captures the localization 'effect.'[12] The final term, ϵ_{iat}, is a residual which I assume to be uncorrelated across area-industry-year observations.

In estimating (1), I limit the sample of area-industry-year observations to those in which there are at least 10 Census observations so that each decile can be identified from a unique observation. This produces samples of 2010 state-industry-years covering each of the 50 states and the District of Columbia, and 4521 city-industry-years encompassing roughly 200 metropolitan areas.[13] Still, because a 90-10 wage difference based on 10 observations likely involves more noise than one based on 1000 observations, I use a generalized/weighted least squares (GLS) technique in which each observation is weighted by the number of Census observations used in the inequality calculations.[14]

Results from two different specifications of equation (1) appear in Table 1. The first,

[10]Using percentiles of the unconditional earnings distribution to represent the wages of high- and low-skill workers is a common practice in the inequality literature (see Juhn et al. (1993)).

[11]Different areas or industries, for instance, may employ different distributions of workers by skill or utilize different wage-setting procedures which would affect measured wage dispersion.

[12]Note, although I use the term 'effect' throughout the paper for expositional purposes, the estimated coefficients should be interpreted as partial correlations, not as measures of causality. How these partial correlations may relate to the actual causal effects is discussed in Section 3.5.

[13]Numbers of observations by year are reported in the footnotes to Tables A2 and A3.

[14]Results from unweighted regressions, incidentally, were very similar to those reported here.

specification I, is merely a baseline case in which I limit the regressors to log industry employment alone (in addition to industry, time, and state/city effects) in an effort to focus purely on the localization effect. Although such a model is rather sparse, the resulting coefficients are indicative of this paper's general result: inequality and localization are significantly, negatively associated at both the state- and metropolitan area-levels. In fact, the estimated magnitudes are quite sizable. The point estimates, for example, suggest that a 1 standard deviation increase in an industry's local market employment corresponds to an 8 percentage point decrease in the overall 90-10 wage gap among states, a 10 percentage point decrease among cities.[15] These magnitudes represent approximately 30 percent of the cross sectional standard deviation of the 90-10 wage differentials in the sample.

To see the robustness of this result, consider estimates from the second specification, II, which adds to the regressions several variables commonly associated with wage dispersion: the proportion of the local resident population with at least a college degree, the fraction of total employment in manufacturing, the unionization rate, the unemployment rate, the proportion of the population that is foreign-born, and the logarithm of the resident population. Increases in local human capital, for instance, may serve to reduce the earnings gap between skilled and unskilled workers (e.g. Moretti (2004)), as might the local unionization rate (e.g. Fortin and Lemieux (1997)) and the fraction of total employment engaged in manufacturing (e.g. Bernard and Jensen (2000)). On the other hand, a higher rate of unemployment (Blank and Blinder (1986)) or larger fraction of the population that is foreign-born (Topel (1997)) might contribute to greater wage dispersion.[16]

[15] The standard deviation of log industry employment is approximately 1.6 in the state sample, 1.5 in the metropolitan area sample.

[16] Log resident population is added simply to pick up any influence of the overall size of the market.

What the estimated coefficients reveal, interestingly, offers some support for these notions. In particular, the college fraction produces a negative coefficient, suggesting that a larger stock of local human capital tends to reduce the earnings gap between high-skill and low-skill workers, possibly due to diminishing marginal productivity of high-skill labor as their numbers in the local labor force increase. Similarly, the unionization rate generates a negative coefficient, which is consistent with the finding that union wages tend to exhibit less variation than non-union wages.[17] The fraction foreign-born enters positively, which is compatible with the findings discussed by Topel (1997) that local markets in the U.S. experiencing greater immigration have also witnessed more rapid increases in inequality.

The estimated coefficients on the manufacturing and unemployment rates, by contrast, are different than what one might have expected. The extent of manufacturing in the local economy enters positively, whereas the rate of unemployment enters negatively. The more significant of these two variables, the unemployment rate, may simply be reflecting changes in the composition of workers over the business cycle: unemployment rises as workers from the bottom end of the wage distribution are laid off, producing a distribution with lower variance (e.g. Solon et al. (1994)).

More importantly, even though several of these regressors enter significantly, the coefficients on log industry employment are virtually unchanged, either in magnitude or statistical significance, when compared to specification I. This feature of the results suggests that, although factors such as unionization and local human capital may be important elements

[17] Some might note that, while this is true for union employees, greater unionization may actually increase wage inequality across both union and non-union workers together (possibly because union workers receive a sizable wage premium). However, the finding that, on net, the effect of unionization on wage inequality is negative is consistent with previous findings (e.g. Bernard and Jensen (2000)).

in understanding the degree of wage dispersion exhibited in a state/city-industry, the size of the state/city-industry itself is an important element too.[18]

Accordingly, while previous work has established that average wage earnings taken across all workers within a local market industry tend to rise as total employment in that industry increases, these findings demonstrate that wage increases at the bottom end of the distribution are, on average, larger than those at the top.[19] Whatever mechanisms underlie localization economies, therefore, also have an equalizing effect on labor earnings.

3.2 Between-Education-Group Inequality

As already noted, the preceding analysis categorizes the skill levels of workers by their places in the distribution of earnings: the 90th percentile represents a high-skill worker, the 10th percentile represents a low-skill worker. This section considers an alternative grouping scheme based on the average earnings of workers within five educational attainment categories: no high school (0 to 8 years of schooling completed), some high school (9 to 11 years), high school (12 years), some college (13 to 15 years), and college or more (16 or more years). After computing the average weekly wage across all workers within the same educational attainment group and same area-industry, I construct four measures of inequality by taking the differences between the average wage for college graduates and

[18] I also estimated (1) with a set of industry-time dummies added to the regressors to account for industry-specific trends in inequality, say due to changes in technology or international trade patterns over time. Doing so, however, did not substantially alter the results.

[19] Estimated localization effects based on regressions of state-industry percentiles on log industry employment (along with industry, area, and time effects) produce coefficients (standard errors) of 0.022 (0.005) for the 90th, 0.069 (0.006) for the 10th. Among cities, they are 0.023 (0.004) for the 90th, 0.092 (0.005) for the 10th.

that of each of the remaining four categories. Doing so generates four different measures of an area-industry's 'high-skill/low-skill' wage gap. These measures are then used as the dependent variable in the estimation of (1) instead of the overall 90-10 differential.

In this case, I maintain the restriction that all area-industries used in the estimation have at least 10 observations from the Census samples. However, I now weight observations by the minimum number of education-group observations associated with a particular inequality measure. That is, if a state- or city-industry has 10 college-graduate observations and 15 high school-graduate observations, its college-high school wage gap is given weight equal to 10. The resulting estimates for the same two specifications described in the last section appear in Table 2.

For the most part, they reveal many of the same qualitative conclusions drawn above for overall inequality. Higher unemployment rates and fractions of foreign-born residents in a state or city, for example, tend to be positively associated with between-education-group gaps, whereas unionization rates are negatively associated with these gaps. Human capital, as captured by the college fraction, also tends to be negatively associated with the wage differences between college educated workers and their less-educated counterparts, particularly at the metropolitan area level.

The most consistently significant regressor of all of those considered, however, is total industry employment. As with overall inequality, earnings gaps between workers belonging to different educational attainment groups tend to decrease significantly as an industry's overall scale within a state or city rises. The point estimates, which are very similar across both specifications, suggest modest, but economically important associations. A 1 standard deviation increase in state-industry employment, for instance, is associated with a 2.5 to

3 percentage point decrease in the wage gap between college graduates and high school graduates. Within metropolitan areas, a 1 standard deviation increase is associated with a 3.5 percentage point decrease in this gap. These magnitudes represent approximately 7 to 10 percent of the standard deviation of the college-high school gaps over the sample period. Similar, albeit somewhat smaller, implied associations hold for the college-some college and college-some high school gaps as well.

3.3 Residual Inequality

The results thus far indicate that the overall distribution of wages within state- and city-industries becomes less disperse as the scale of those state- or city-industry grows larger. Moreover, they indicate that part of this decrease in dispersion operates in a between-education-group channel: the gaps between workers of different education groups grow narrower as area-industry employment increases. Yet, since overall inequality consists of both between-group inequality and within-group inequality, do we also see earnings dispersion among workers with the *same* levels of education decrease with localization too?

To answer this question, I turn to the investigation of 'residual' inequality which measures the degree of wage variation among workers with the same observable characteristics. The analysis proceeds in two stages. In the first, I estimate a standard Mincerian wage regression in which the weekly earnings of worker j employed in industry i of area a in year t, w^j_{iat}, is specified as

$$\log(w^j_{iat}) = \alpha_{at} + \beta_{at} X^j_{iat} + v^j_{iat} \tag{2}$$

where X^j_{iat} is a vector of personal characteristics for this worker, including a quartic in potential experience[20], four educational attainment dummies (some high school, high school, some college, college or more), and a marital status indicator. Notice, because returns to these observable features likely vary across both local markets (e.g. Dahl (2002)) and time (e.g. Juhn et al. (1993)), this equation is estimated separately for each state or city in each year.[21] After estimating (2), I take the 90th and 10th percentiles of the fitted residuals within each area-industry-year to form a high-skill/low-skill residual wage gap, $(\hat{v}^H_{iat} - \hat{v}^L_{iat})$, which is then modeled in a manner analogous to the overall and between-education-group wage differentials:

$$\left(\hat{v}^H_{iat} - \hat{v}^L_{iat}\right) = \delta_i + \delta_a + \delta_t + \theta Z_{at} + \gamma \log(N_{iat}) + \epsilon_{iat} \tag{3}$$

Results appear in Table 3.[22]

On the whole, they provide similar conclusions to those already drawn. The local unionization rate, again, generates a significantly negative coefficient at both levels of geographic aggregation, while there is also some evidence that inequality is influenced negatively by the unemployment rate among states, positively by the percentage of total employment in manufacturing among cities.

[20]Potential experience is calculated as the maximum of (age-years of education-6) and 0. Education in 1990 is estimated using Table 5 of Park (1994).

[21]Although wage regressions like (2) also often include industry dummies to account for inter-industry wage differentials (e.g. Bartel and Sicherman (1999)), doing so is unnecessary in this case because I am investigating wage differentials within area-industry-years.

[22]As above, adding industry-time effects to the regression did not alter the estimated localization coefficients substantially.

As for the estimated localization effects on inequality, all are significantly negative and suggestive of reasonably large magnitudes. Point estimates, for example, suggest that a 1 standard deviation increase in an industry's employment is accompanied by a 5 to 6.6 percentage point decline in the residual wage difference. These magnitudes are approximately 25 percent of the standard deviation of the residual 90-10 wage differentials in the sample. Clearly, such findings indicate that localization is indeed associated with reduced earnings disparity among workers with similar observable characteristics, including educational attainment.

Moreover, the fact that these point estimates are rather large when compared to the estimates for overall inequality in Table 1 suggests that the majority of the decrease in the overall 90-10 wage gap with localization involves decreases in residual inequality rather than decreases in the gaps between workers with different observable measures of skill. Hence, as local employment expands, we tend observe a decrease in overall wage dispersion largely because workers within the same education and experience groups exhibit less wage variation.[23] This conclusion is also consistent with the evidence reported in the previous section which indicates that, although negative, the association between localization and the wage gaps between workers of different education groups is rather modest.

3.4 Industry-Specific Results

Because localization effects on average productivity have been shown to differ in magnitude across manufacturing sectors (e.g. Henderson (1986)), there may also be inter-industry

[23] This basic result is compatible with the findings of Juhn et al. (1993) who find that changes in residual inequality were a major component of changes in overall inequality for workers in the U.S. over this time period.

differences in localization's association with inequality. To account for this possibility, I repeat the analysis above allowing the coefficients on log industry employment to vary by two-digit sector. Results, which for the sake of conciseness are limited to the longer specification, II, appear in Tables 4A (for states) and 4B (for metropolitan areas).

Most noticeably, the vast majority of the estimates are negative, suggesting that the pooled results documented in Tables 1-3 hold qualitatively for a variety of different industries within the manufacturing sector. This feature is most strikingly evident when considering the overall and residual 90-10 wage gaps, where 75 of the 76 coefficients across the two levels of geographic aggregation are negative.[24] Moreover, many of these negative associations are important statistically. Among states, 17 industries produce significantly negative localization coefficients when considering overall inequality, 14 when considering residual inequality. Among metropolitan areas, 15 industries exhibit statistically important negative associations between city-level employment and overall inequality whereas 16 do so for residual inequality.

The estimated localization coefficients for the between-education-group gaps also turn out to be predominantly negative although there is greater heterogeneity in this regard than what was observed for the overall and residual measures. Of the 20 coefficients reported for each between-education-group gap at the state level, for example, 10 are negative for the college-no high school gap (6 significant), 15 for the college-some high school gap (7 significant), 15 for the college-high school gap (10 significant), 16 for the college-some college gap (9 significant). Among metropolitan areas, the numbers of negative coefficients (and

[24] The Census samples did not produce any cities with at least 10 white males (meeting the above-stated criteria) for either industry 24 (Lumber and Wood Products) or 25 (Furniture and Fixtures).

the number of these that differ statistically from zero) are 12 (3) for the college-no high school gap, 14 (9) for the college-some high school gap, 16 (11) for the college-high school gap, and 16 (8) for the college-some college gap.

To be sure, within each measure of inequality, the results do show some heterogeneity from one industry to another. Among the state sample, for instance, the estimated localization coefficient ranges between -0.01 (SIC 23 - Apparel and Other Textile Products) and -0.17 (SIC 24 - Lumber and Wood Products) for the overall measure; 0.03 (SIC 25 - Furniture and Fixtures and SIC 30 - Rubber and Miscellaneous Plastics Products) and -0.07 (SIC 24 - Lumber and Wood Products) for the college-high school gap; 0.01 (SIC 23 - Apparel and Other Textile Products) and -0.11 (SIC 24 - Lumber and Wood Products) for the residual. Within the sample of cities, the coefficients show somewhat less variation, extending from -0.014 (SIC 22 - Textile Mill Products) to -0.09 (SIC 37 - Transportation Equipment) for overall inequality; 0.017 (SIC 30 - Rubber and Miscellaneous Plastics Products) to -0.046 (SIC 33 - Primary Metal Industries) for the college-high school gap; -0.01 (SIC 23 - Apparel and Other Textile Products and SIC 39 - Miscellaneous Industries) to -0.06 (SIC 37) for the residual measure.

Given that there is some overlap in the industries comprising the extreme values of these localization effects (e.g. SIC 23), one might surmise that the estimates in Tables 4A and 4B are positively associated across inequality measures. This hypothesis turns out to be true. The correlation between the localization coefficients estimated for overall inequality and those estimated for residual inequality, for example, is 0.87 among states, 0.64 among cities. Similarly, the overall inequality coefficients and the college-high school coefficients show a correlation of 0.61 among states, 0.64 among cities. Hence, industries which exhibit a strong inverse association between localization and one particular inequality measure also tend to exhibit a strong inverse association with the other measures too.

More generally, however, this evidence shows that there is remarkable consistency across different manufacturing sectors in the association between localization and inequality. Hence, just as previous work has indicated that a broad collection of industries seem to experience significant gains in their average or aggregate productivity as their employment within a local market rises, these results suggest a similar degree of uniformity for localization's qualitative impact on inequality.

3.5 Endogeneity Considerations

The analysis so far has treated the regressors in equations (1) and (3) as exogenous. However, there is ample reason to suspect that at least two of them – total employment within a state- or city-industry and the overall fraction of college-educated individuals in the local market – may themselves be determined by the level of inequality. Were this the case, the results established thus far would be biased.

Of course, a priori, it is uncertain as to whether the sign of any such bias is positive or negative. Higher inequality might, for instance, be associated with decreases in an industry's employment, say because mobile firms and workers view it negatively and, therefore, choose to locate elsewhere. Alternatively, an increase in inequality may generate an increase in employment, possibly because high inequality is associated with low 'bottom-end' wages which attract producers seeking cheap labor.

Similarly, high inequality may either serve to raise or lower the college fraction depending on how it influences the location decisions of workers with varying levels of education. High inequality may, for example, be associated with particularly high wages for skilled workers, and thus serve to attract greater numbers of college-educated individuals. It also may be associated with especially low wages for less-educated workers which should provide an incentive for these workers to leave. Both of these possibilities, naturally, should increase the college fraction. On the other hand, if viewed negatively, high inequality may lead to decreases in the college fraction as highly educated workers (who tend to be more mobile than less-educated workers (e.g. Dahl (2002)) locate elsewhere.

In an effort to get a handle on these issues, I consider the following statistical exercise. Using the samples of state- and city-industries, I estimate two sets of regressions: one in which the growth of a state- or city-industry's employment is expressed as a function of the initial level of inequality (overall, residual, between-education-group), and one in which the change in a local market's college fraction is a function of initial inequality.[25] Results appear in Table 5.

Beginning with the coefficients from the employment growth regressions, it is evident that two of the inequality measures - the overall and residual 90-10 wage differences - are strongly associated with future rates of growth at both the state and city level. All else equal, state- and city-industries with higher levels of inequality defined by these two measures grow faster than those with lower levels. The same pattern does not hold, however,

[25] Each regression also contains three region dummies (West, Midwest, Northeast) to account for any exogenous, geographic variation in the dependent variables. The employment growth equations also control for the initial college fraction and manufacturing rate since these two variables have been shown to influence local market growth (e.g. Glaeser et al. (1995)). Dropping these two variables produced similar estimates. Finally, because the state-level regressions involve growth rates for two different decades (1970-80, 1980-90), I further include a time dummy.

when inequality is defined in a between-education-group manner. The resulting coefficients are insignificant in each of these instances. Such findings, therefore, only offer some limited support for the notion that state- and city-industry growth is driven by low labor costs, particularly at the bottom end of the earnings distribution.

Turning to the estimates from the college fraction regressions, one can see that the majority of the coefficients across all of the inequality measures are significantly positive. Only 1 of the 12 coefficients in the second column of results in Table 5 is negative (and insignificant). Of the remaining 11 positive coefficients, 9 differ statistically from zero indicating that, on average, states and cities with greater inequality among their resident industries do tend to experience greater human capital growth. As suggested above, this result may reflect the movement of highly educated workers into markets where the relative returns paid to them are high.

On the whole, these findings suggest that these two regressors - log industry employment and the college fraction - are indeed endogenous with respect to inequality. Hence, the parameter estimates reported thus far are likely biased. However, the results in Table 5 also suggest that the sign of the bias may actually lead the coefficients reported in Tables 1-4 to *understate* the magnitude of the inequality-localization association.

Consider, for example, a stochastic shock that increases a state- or city-industry's level of inequality. Following the estimates in Table 5, this positive shock would tend to increase employment growth and a state or city's college fraction subsequently. Recall, all of the statistically significant coefficients reported in Table 5 are positive. As a consequence, there is likely a positive association between the error term in the inequality regressions and each of these regressors (if there is any association at all). This correlation would then bias the estimated coefficients on these two regressors upward which, in this case, is in the direction of zero. One should, therefore, view the associations reported in Tables 1-4 as *understating* the true causal effect of localization on inequality. The fact that the GLS results indicate negative (and largely significant) associations between inequality and each of these regressors in spite of this bias only reinforces the conclusions drawn above.

4 Concluding Comments

This paper has explored the nature of localization economies by examining how wage dispersion varies with the magnitude of an industry's presence in a particular geographic market. The results are unambiguous: an increase in an industry's employment within a state or metropolitan area is accompanied by significant decreases, on average, in wage inequality. What is more, this relationship holds for a variety of inequality measures and is robust to the inclusion of controls for a number of inequality-related features, including rates of union membership, the presence of foreign-born labor, and levels of education.

While certainly interesting in itself, this evidence may also enhance our understanding of two larger issues. First, it may offer some additional insight into the rise of earnings inequality in the U.S., which has spawned a massive literature over the past two decades. In particular, although the decline in manufacturing activity has been identified as a potential determinant of the rise in inequality (e.g. Bernard and Jensen (2000)), the results reported here suggest that there may be a geographic aspect to this relationship.

As demonstrated in Tables A2 and A3 of the Appendix, there has been a decrease in the average state- and city-industry employment for a typical manufacturing worker over the past several decades. Hence, part of the rise in earnings disparity in the U.S.

over the past 30 years (at least, among manufacturing workers) may be the product of smaller (local) concentrations of industry and not simply the decline of manufacturing at the national level. Such a conjecture would also imply that decreases in an aggregate economy's manufacturing activity that are accompanied by *growing* geographic concentration - say, given by the consolidation of several different industrial clusters into a single cluster - would be associated with *lower* inequality. Although not a straightforward task empirically, evaluating this implication would certainly provide an interesting topic for future work.

Second, these results provide further evidence on the nature of localization economies which, as noted in the Introduction, has also attracted a very large literature. Unfortunately, our understanding of why the geographic concentration of industry is positively associated with productivity remains limited. To be sure, a number of theories have been advanced over the last century, most notably Marshall's (1920) famous three, which hold that productivity gains stem from (*i*) the local spillover of industry specific knowledge, (*ii*) the creation of an extensive array of specialized input providers that allows producers to benefit from a more extensive division of labor, and (*iii*) improved firm-worker matching by making labor market search easier. Yet, only recently has there been any attempt to test these theories directly.[26]

The evidence documented in this paper, of course, does not provide a direct evaluation of these ideas. However, it does offer a more detailed description of the productivity gains tied to localization than do studies of average or aggregate productivity. Again, the results suggest that localization tends to to have a larger effect on the productivity of low-skill workers than on the productivity of high-skill workers. Rationalizing the extent to which

[26]Two prominent examples are Dumais et al. (1997) and Rosenthal and Strange (2001).

Marshall's (1920) three theories (or any others) are consistent with this pattern would also be an interesting avenue for future research.

Table 1: Localization and Overall Inequality

	Specification			
Variable	*I*	*II*	*I*	*II*
Log Industry Employment	-0.048^c	-0.049^c	-0.069^c	-0.069^c
	(0.006)	(0.005)	(0.004)	(0.004)
College Rate	–	-0.99^b	–	-0.8
		(0.48)		(0.54)
Manufacturing Rate	–	0.39	–	0.24
		(0.25)		(0.31)
Union Rate	–	-0.5^b	–	-0.48^b
		(0.2)		(0.22)
Unemployment Rate	–	-1.06^b	–	-0.11
		(0.48)		(0.5)
Foreign Rate	–	1.02^c	–	0.41
		(0.39)		(0.36)
Log Population	–	-0.05	–	-0.08
		(0.07)		(0.1)
R^2	0.78	0.79	0.7	0.7
Sample	States	States	Cities	Cities

Note: GLS estimates. Dependent variable is 90-10 difference in log weekly wages. Each regression also includes industry, state/city, and time effects. Regressions are weighted by the number of individual observations used in the percentile calculations for each state/city-industry-year. Heteroskedasticity-consistent standard errors are reported in parentheses. Superscript *a* denotes significance at 10 percent, *b* at 5 percent; *c* at 1 percent. 2010 state-industry-year observations, 4521 city-industry-year observations.

Table 2: Localization and Between-Education-Group Inequality

Dep. Variable	Spec.	Log Ind. Emp.	College Rate	Manuf. Rate	Union Rate	Unemp. Rate	Foreign Rate	Log Pop.	Sample
College-No HS	I	-0.007	–	–	–	–	–	–	States
		(0.006)							
	II	-0.008	0.8	1.5^c	-0.9^c	1.1^b	2.5^c	-0.1	States
		(0.006)	(0.6)	(0.3)	(0.2)	(0.52)	(0.4)	(0.07)	
College-Some HS	I	-0.012^b	–	–	–	–	–	–	States
		(0.005)							
	II	-0.015^c	-0.36	0.99^c	-0.33^a	0.69	1.3^c	-0.02	States
		(0.005)	(0.5)	(0.27)	(0.19)	(0.53)	(0.35)	(0.06)	
College-HS	I	-0.015^c	–	–	–	–	–	–	States
		(0.005)							
	II	-0.019^c	-1.05^b	0.59^b	-0.21	0.21	0.86^c	0.01	States
		(0.005)	(0.47)	(0.25)	(0.18)	(0.47)	(0.3)	(0.06)	
College-Some Coll.	I	-0.012^c	–	–	–	–	–	–	States
		(0.004)							
	II	-0.015^c	-0.68^a	0.3	0.05	0.05	0.2	0.007	States
		(0.004)	(0.4)	(0.2)	(0.1)	(0.4)	(0.2)	(0.05)	
College-No HS	I	-0.012^b	–	–	–	–	–	–	Cities
		(0.006)							
	II	-0.013^b	-0.26	0.38	-0.5	2.2^c	0.26	0.18	Cities
		(0.006)	(0.62)	(0.41)	(0.33)	(0.71)	(0.35)	(0.13)	
College-Some HS	I	-0.018^c	–	–	–	–	–	–	Cities
		(0.005)							
	II	-0.019^c	-1.6^c	-0.005	0.003	0.96	0.75^b	0.1	Cities
		(0.005)	(0.54)	(0.33)	(0.3)	(0.6)	(0.3)	(0.1)	
College-HS	I	-0.024^c	–	–	–	–	–	–	Cities
		(0.004)							
	II	-0.024^c	-1.8^c	-0.05	-0.19	0.79	0.68^c	-0.007	Cities
		(0.004)	(0.45)	(0.3)	(0.2)	0.5)	(0.26)	(0.08)	
College-Some Coll.	I	-0.015^c	–	–	–	–	–	–	Cities
		(0.004)							
	II	-0.016^c	-1.07^c	0.26	-0.23	0.73^a	-0.07	0.04	Cities
		(0.004)	(0.4)	(0.25)	(0.18)	(0.43)	(0.2)	(0.07)	

Note: GLS estimates. Each regression also includes industry, state/city, and time effects. Heteroskedasticity-consistent standard errors are reported in parentheses. Numbers of state-industry-year observations are 1886 for college-no high school, 1937 for college-some high school, 1970 for college-high school, 1964 for college-some college. Numbers of metropolitan area-industry-year observations are 3536 for college-no high school, 4101 for college-some high school, 4375 for college-high school, 4328 for college-some college. Superscript a denotes significance at 10 percent, b at 5 percent; c at 1 percent.

Table 3: Localization and Residual Inequality

	Specification			
Variable	*I*	*II*	*I*	*II*
Log Industry Employment	-0.032^c	-0.033^c	-0.044^c	-0.044^c
	(0.004)	(0.004)	(0.003)	(0.003)
College Rate	–	-0.13	–	-0.07
		(0.32)		(0.36)
Manufacturing Rate	–	0.25	–	0.4^a
		(0.16)		(0.21)
Union Rate	–	-0.43^c	–	-0.23^a
		(0.16)		(0.14)
Unemployment Rate	–	-0.69^b	–	-0.07
		(0.28)		(0.32)
Foreign Rate	–	0.45	–	0.04
		(0.3)		(0.23)
Log Population	–	0.04	–	-0.07
		(0.06)		(0.07)
R^2	0.78	0.78	0.67	0.67
Sample	States	States	Cities	Cities

Note: GLS estimates. Dependent variable is 90-10 difference in residual log weekly wages. Each regression also includes industry, state/city, and time effects. Regressions are weighted by the number of individual observations used in the percentile calculations for each state/city-industry-year. Heteroskedasticity-consistent standard errors are reported in parentheses. Superscript a denotes significance at 10 percent, b at 5 percent; c at 1 percent. 2010 state-industry-year observations, 4521 city-industry-year observations.

Table 4A: Localization and Inequality - States

By Two-Digit Industry

SIC	Overall 90-10	Residual 90-10	College-No HS	College-Some HS	College-HS	College-Some Coll.
20	-0.08^c	-0.043^c	-0.05^c	-0.05^b	-0.05^c	-0.04^c
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.016)
21	-0.1^c	-0.03	-0.11^c	-0.1^b	-0.03	-0.04
	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
22	-0.06^c	-0.05^c	-0.03^c	-0.02^b	0.002	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
23	-0.01	0.01	0.02	0.03	0.001	-0.03
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
24	-0.17^c	-0.11^c	-0.13^a	-0.1	-0.07	0.05
	(0.03)	(0.03)	(0.07)	(0.06)	(0.07)	(0.06)
25	-0.05^b	-0.01	0.05	0.05	0.03	-0.0004
	(0.02)	(0.01)	(0.06)	(0.06)	(0.05)	(0.06)
26	-0.06^c	-0.04^c	0.005	0.009	0.007	0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
27	-0.034^b	-0.01	0.005	-0.02^a	-0.026^b	-0.018^a
	(0.015)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
28	-0.065^c	-0.04^c	-0.01	-0.016	-0.03^b	-0.03^a
	(0.02)	(0.01)	(0.01)	(0.2)	(0.015)	(0.016)
29	-0.07^c	-0.05^c	-0.02	-0.03	-0.017	-0.01
	(0.02)	(0.01)	(0.02)	(0.02)	(0.015)	(0.01)
30	-0.017	-0.02^b	0.05^b	0.02	0.03	0.04^b
	(0.018)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)
31	-0.026^a	-0.01	-0.01	-0.015	-0.015	-0.007
	(0.015)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
32	-0.07^c	-0.03^c	0.005	0.003	-0.02	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
33	-0.043^c	-0.022^c	-0.025^b	-0.03^b	-0.03^b	-0.015
	(0.01)	(0.006)	(0.01)	(0.012)	(0.01)	(0.01)
34	-0.06^c	-0.03^c	-0.01	-0.03^c	-0.04^c	-0.03^c
	(0.01)	(0.006)	(0.01)	(0.01)	(0.01)	(0.009)
35	-0.06^c	-0.033^c	-0.03^c	-0.03^c	-0.04^c	-0.03^c
	(0.01)	(0.006)	(0.01)	(0.01)	(0.01)	(0.007)
36	-0.044^c	-0.03^c	0.006	-0.008	-0.02^b	-0.02^c
	(0.012)	(0.007)	(0.01)	(0.01)	(0.01)	(0.007)
37	-0.066^c	-0.05^c	0.01	-0.01	-0.016^b	-0.015^b
	(0.01)	(0.006)	(0.01)	(0.01)	(0.008)	(0.007)
38	-0.019	-0.02^b	0.01	-0.007	-0.02^b	-0.02^c
	(0.012)	(0.01)	(0.01)	(0.01)	(0.01)	(0.006)
39	-0.047^c	-0.006	0.002	-0.02	-0.03^b	-0.02
	(0.015)	(0.02)	(0.02)	(0.02)	(0.016)	(0.02)

Note: GLS estimates of coefficients on log industry employment. Dependent variable is inequality. Regressions follow specification II described in Tables 1-3. Heteroskedasticity-consistent standard errors are reported in parentheses. Superscript a denotes significance at 10 percent, b at 5 percent; c at 1 percent.

Table 4B: Localization and Inequality - Cities

By Two-Digit Industry

SIC	Overall 90-10	Residual 90-10	College-No HS	College-Some HS	College-HS	College-Some Coll.
20	-0.08^c	-0.04^c	-0.014	-0.022^a	-0.018	-0.017
	(0.01)	(0.01)	(0.01)	(0.012)	(0.012)	(0.014)
21	-0.03	-0.04^b	-0.11^c	-0.09	-0.013	-0.07^c
	(0.03)	(0.02)	(0.03)	(0.06)	(0.03)	(0.02)
22	-0.01	-0.02^a	-0.015	-0.014	-0.01	-0.023
	(0.02)	(0.01)	(0.014)	(0.012)	(0.01)	(0.015)
23	-0.05^a	-0.01	0.001	-0.01	-0.02	-0.048^c
	(0.03)	(0.01)	(0.01)	(0.016)	(0.02)	(0.02)
24	–	–	–	–	–	–
25	–	–	–	–	–	–
26	-0.04^c	-0.023^c	0.01	0.003	-0.001	-0.003
	(0.01)	(0.006)	(0.01)	(0.01)	(0.01)	(0.01)
27	-0.06^c	-0.026^c	0.008	-0.026^c	-0.03^c	-0.024^c
	(0.01)	(0.006)	(0.01)	(0.01)	(0.008)	(0.008)
28	-0.06^c	-0.05^c	-0.006	-0.004	-0.015^b	-0.003
	(0.01)	(0.005)	(0.009)	(0.01)	(0.007)	(0.008)
29	-0.08^c	-0.05^c	0.014	0.01	0.0003	0.03
	(0.03)	(0.015)	(0.03)	(0.02)	(0.02)	(0.02)
30	-0.016	-0.024^c	0.04^c	0.03^b	0.017	0.013
	(0.01)	(0.008)	(0.01)	(0.01)	(0.013)	(0.013)
31	-0.064^c	-0.034^c	-0.024	-0.01	-0.025^b	-0.014
	(0.02)	(0.01)	(0.015)	(0.01)	(0.011)	(0.01)
32	-0.08^c	-0.04^c	-0.01	-0.05^c	-0.031^c	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.014)
33	-0.1^c	-0.05^c	-0.05^c	-0.046^c	-0.046^c	-0.023^c
	(0.01)	(0.005)	(0.01)	(0.01)	(0.007)	(0.008)
34	-0.07^c	-0.034^c	-0.003	-0.02^b	-0.021^c	-0.01
	(0.01)	(0.005)	(0.01)	(0.01)	(0.008)	(0.007)
35	-0.07^c	-0.04^c	-0.03^c	-0.023^c	-0.03^c	-0.014^c
	(0.01)	(0.005)	(0.01)	(0.009)	(0.006)	(0.006)
36	-0.06^c	-0.05^c	0.001	-0.001	-0.017^b	-0.01^a
	(0.01)	(0.005)	(0.01)	(0.01)	(0.007)	(0.005)
37	-0.09^c	-0.06^c	-0.01	-0.022^c	-0.03^c	-0.017^c
	(0.01)	(0.004)	(0.01)	(0.009)	(0.006)	(0.006)
38	-0.07^c	-0.05^c	-0.003	-0.028^c	-0.027^c	-0.011
	(0.01)	(0.007)	(0.01)	(0.01)	(0.009)	(0.007)
39	-0.05^c	-0.01	-0.012	-0.025^a	-0.034^c	-0.028^b
	(0.01)	(0.01)	(0.015)	(0.014)	(0.012)	(0.011)

Note: GLS estimates of coefficients on log industry employment. Dependent variable is inequality. Regressions follow specification II described in Tables 1-3. Heteroskedasticity-consistent standard errors are reported in parentheses. Superscript a denotes significance at 10 percent, b at 5 percent; c at 1 percent.

Table 5: Initial Inequality and the Change in Industry Employment and Local Human Capital

Variable	Industry-Employment Growth	Change in College Fraction	Sample
Overall 90-10 Differential	0.32^c	0.004^c	States
	(0.07)	(0.001)	
	0.17^c	0.006^c	Cities
	(0.05)	(0.001)	
Residual 90-10 Differential	0.33^c	-0.001	States
	(0.09)	(0.001)	
	0.13^b	0.004^c	Cities
	(0.06)	(0.001)	
College-No HS Gap	0.06	0.004^c	States
	(0.07)	(0.001)	
	-0.05	0.002^b	Cities
	(0.04)	(0.001)	
College-Some HS Gap	0.04	0.004^c	States
	(0.06)	(0.001)	
	-0.007	0.001	Cities
	(0.04)	(0.0006)	
College-HS Gap	-0.009	0.004^c	States
	(0.08)	(0.001)	
	0.04	0.002^c	Cities
	(0.04)	(0.001)	
College-Some College Gap	-0.02	0.003^c	States
	(0.07)	(0.001)	
	-0.06	0.001	Cities
	(0.05)	(0.001)	

Note: Coefficients on initial levels of inequality. Dependent variables are the changes in state/city-industry log employment and the change in state/city-level college fraction. Initial fractions of total employment in manufacturing and the initial college fraction are included in the industry-employment growth regressions. All specifications have three geographic dummies – West, Midwest, and Northeast regions – and the state-level regressions contain a time dummy for the 1970-80 period. Heteroskedasticity-consistent standard errors are reported in parentheses. Superscript a denotes significance at 10 percent, b at 5 percent; c at 1 percent.

Appendix

Table A1: Industries and Codes

SIC	Industry Name	1970 IPUMS Industry Codes	1980/90 IPUMS Industry Codes
20	Food and Kindred	268-298	100-122
21	Tobacco	299	130
22	Textile Mill Products	307-318	132-150
23	Apparel and Other Textile Products	319-327	151-152
24	Lumber and Wood Products	107-109	230-241
25	Furniture and Fixtures	118	242
26	Paper and Allied Products	328-337	160-162
27	Printing and Publishing	338-339	171-172
28	Chemicals and Allied Products	347-369	180-192
29	Petroleum and Coal Products	377-378	200-201
30	Rubber and Miscellaneous Plastics Products	379-387	210-212
31	Leather and Leather Produts	388-397	220-222
32	Stone, Clay, Glass, and Concrete Products	119-138	250-262
33	Primary Metal Industries	139-149	270-280
34	Fabricated Metal Products	157-168	281-300
35	Industrial Machinery and Equipment	177-198	310-332
36	Electrical and Electronic Equipment	199-209	340-350
37	Transportation Equipment	219-238, 258	351-370
38	Instruments and Related Products	239-257	371-390
39	Miscellaneous Industries	259	391

Table A2: Summary Statistics – States

Year	Variable	Mean	Standard Deviation	Minimum	Maximum
1970	90-10 Overall Wage Difference	1.12	0.24	0.33	2.92
	90-10 Residual Wage Difference	0.91	0.2	0.44	2.91
	College-No HS Gap	0.66	0.31	-0.7	2.4
	College-Some HS Gap	0.6	0.31	-0.79	2.5
	College-HS Gap	0.52	0.28	-0.62	2.34
	College-Some College Gap	0.44	0.3	-0.92	2.11
	State-Industry Employment	26082.8	44667.6	144	345960
	College Rate	0.11	0.02	0.067	0.18
	Manufacturing Rate	0.22	0.1	0.047	0.36
	Union Rate	0.25	0.08	0.088	0.42
	Unemployment Rate	0.045	0.01	0.027	0.092
	Foreign Rate	0.033	0.025	0.006	0.11
	Population	3984566	4314048	300382	19957715
	Census Observations Per State-Industry	552.8	1092.2	10	9003
1980	90-10 Overall Wage Difference	1.24	0.25	0.53	3.15
	90-10 Residual Wage Difference	0.98	0.18	0.52	2.12
	College-No HS Gap	0.54	0.29	-1.6	2.1
	College-Some HS Gap	0.53	0.28	-1.6	1.85
	College-HS Gap	0.43	0.23	-1.6	1.44
	College-Some College Gap	0.33	0.24	-1.54	1.5
	State-Industry Employment	23417.7	37860.6	24	317230
	College Rate	0.16	0.03	0.1	0.27
	Manufacturing Rate	0.2	0.08	0.045	0.33
	Union Rate	0.21	0.075	0.06	0.35
	Unemployment Rate	0.064	0.015	0.037	0.11
	Foreign Rate	0.043	0.034	0.01	0.14
	Population	4442075	4699160	401851	23667902
	Census Observations Per State-Industry	804.8	1435.4	10	18334
1990	90-10 Overall Wage Difference	1.39	0.28	0.55	3.48
	90-10 Residual Wage Difference	1.1	0.21	0.43	2.28
	College-No HS Gap	0.71	0.36	-0.93	2.6
	College-Some HS Gap	0.71	0.31	-0.83	2.22
	College-HS Gap	0.55	0.25	-0.74	1.8
	College-Some College Gap	0.41	0.24	-1.18	1.7
	State-Industry Employment	20923	32242.1	13	306417
	College Rate	0.2	0.04	0.12	0.33
	Manufacturing Rate	0.16	0.06	0.04	0.27
	Union Rate	0.15	0.06	0.046	0.294
	Unemployment Rate	0.06	0.014	0.035	0.096
	Foreign Rate	0.054	0.046	0.01	0.21
	Population	4876664	5439195	453588	29760021
	Census Observations Per State-Industry	725.8	1190.5	10	12534

Note: Unweighted state-industry statistics. Unweighted state-level characteristics (college, manufacturing, unionization, unemployment, and foreign-born rates; resident population) based on the 50 states and the District of Columbia.

Table A3: Summary Statistics – Cities

Year	Variable	Mean	Standard Deviation	Minimum	Maximum
1980	90-10 Overall Wage Difference	1.2	0.31	0.33	2.78
	90-10 Residual Wage Difference	0.94	0.23	0.3	2.35
	College-No HS Gap	0.5	0.39	-2.1	2.3
	College-Some HS Gap	0.5	0.39	-1.1	2.4
	College-HS Gap	0.41	0.34	-1.66	2.25
	College-Some College Gap	0.32	0.36	-1.75	2.2
	City-Industry Employment	6434.8	15992.3	10	210607
	College Rate	0.16	0.05	0.08	0.35
	Manufacturing Rate	0.22	0.09	0.03	0.52
	Union Rate	0.21	0.08	0.06	0.35
	Unemployment Rate	0.07	0.02	0.02	0.15
	Foreign Rate	0.045	0.05	0.002	0.3
	Population	831451.2	1770177	100376	17260490
	Census Observations Per City-Industry	190.4	532.3	10	12932
1990	90-10 Overall Wage Difference	1.35	0.35	0.24	3.7
	90-10 Residual Wage Difference	1.06	0.27	0.39	3.03
	College-No HS Gap	0.7	0.49	-1.8	3.2
	College-Some HS Gap	0.68	0.45	-2.7	2.5
	College-HS Gap	0.54	0.39	-1.46	2.52
	College-Some College Gap	0.41	0.38	-1.99	2.96
	City-Industry Employment	5660.4	13521.6	10	223972
	College Rate	0.2	0.06	0.1	0.37
	Manufacturing Rate	0.17	0.07	0.036	0.46
	Union Rate	0.15	0.07	0.046	0.29
	Unemployment Rate	0.06	0.02	0.03	0.14
	Foreign Rate	0.06	0.07	0.003	0.4
	Population	947122	1971706	106470	17830586
	Census Observations Per City-Industry	154.8	410.7	10	8860

Note: Unweighted city-industry statistics. Unweighted city-level characteristics (college, manufacturing, unionization, unemployment, and foreign-born rates; resident population) based on 203 metropolitan areas for 1980, 198 for 1990.

References

Abdel-Rahman, H. and M. Fujita. (1990) "Product Variety, Marshallian Externalities, and City Sizes." *Journal of Regional Science*. 30, 165-183.

Acemoglu, D. and J. Angrist. (1999) "How Large are the Social Returns to Education? Evidence From Compulsory Schooling Laws." NBER Working Paper No. 7444.

Autor, D., L. Katz, and A. Krueger. (1998) "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*. 113, 1169-1213.

Bartel, A. and N. Sicherman. (1999) "Technological Change and Wages: An Inter-Industry Analysis." *Journal of Political Economy*. 107 (2), 285-325.

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.

Black, D. and V. Henderson. (1999) "A Theory of Urban Growth." *Journal of Political Economy*. 107, 252-284.

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.

Carlino, G. and R. Voith. (1992) "Accounting for Differences in Aggregate State Productivity." *Regional Science and Urban Economics*. 22, 597-617.

Ciccone, A. and R. Hall. (1996) "Productivity and the Density of Economic Activity." *American Economic Review*. 86 (1), 54-70.

Dahl, G. (2002) "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica*. 70 (6), 2367-2420.

Dumais, G., G. Ellison, and E. Glaeser. (1997) "Geographic Concentration as a Dynamic Process." NBER Working Paper No. 6270.

Ellison, G. and E. Glaeser. (1997) "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy*. 105, 889-927.

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

Fuchs, V. (1962) *Changes in the Location of Manufacturing in the United States Since 1929*. New Haven, CT: Yale University Press.

Glaeser, E., J. Scheinkman, and A. Shleifer. (1995) "Economic Growth in a Cross-Section of Cities." *Journal of Monetary Economics*. 36, 117-143.

Henderson, V. (1986) "The Efficiency of Resource Usage and City Size." *Journal of Urban Economics*. 19, 47-70.

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

Hoover, E. (1948) *The Location of Economic Activity*. New York: McGraw-Hill.

Juhn, C., K. Murphy, and B. Pierce. (1993) "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy.* 101, 410-442.

Katz, L. and K. Murphy. (1992) "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics.* 107, 35-78.

Krugman, P. (1991) *Geography and Trade.* Cambridge, MA: MIT Press.

Marshall, A. (1920) *Principles of Economics.* London: Macmillan.

Moomaw, R. (1981) "Productivity and City Size: A Critique of the Evidence." *Quarterly Journal of Economics.* 96, 675-688.

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

Park, J. (1994) "Estimation of Sheepskin Effects and Returns to Schooling Using the Old and New CPS Measures of Educational Attainment." Princeton Industrial Relations Section Working Paper No. 338.

Perloff, H., E. Dunn, E. Lampard, and R. Muth. (1960) *Regions, Resources, and Economic Growth.* Baltimore, MD: Johns Hopkins Press.

Pred, A. (1966) *The Spatial Dynamics of U.S. Urban-Industrial Growth, 1800-1914.* Cambridge, MA: MIT Press.

Rosenthal, S. and W. Strange. (2001) "The Determinants of Agglomeration." *Journal of Urban Economics.* 50, 191-229.

Ruggles, S. and M. Sobek et al. (2003) *Integrated Public Use Microdata Series: Version 3.0.* Minneapolis: Historical Census Projects, University of Minnesota.

Solon, G., R. Barsky, and J. Parker. (1994) "Measuring the Cyclicality of Real Wages: How Important is Composition Bias." *Quarterly Journal of Economics.* 109, 1-25.

Sveikauskas, L. (1975) "The Productivity of Cities." *Quarterly Journal of Economics.* 89, 392-413.

Topel, R. (1997) "Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality." *Journal of Economic Perspectives.* 11, 55-74.

U.S. Bureau of the Census. (1971) *County Business Patterns, 1970.* Washington, D.C.: The Bureau.

U.S. Bureau of the Census. (1974) *County and City Data Book, 1972* [computer file]. ICPSR, ed. Ann Arbor, MI: ICPSR [producer and distributor].

U.S. Bureau of the Census. (1982) *County Business Patterns, 1980* [machine-readable data file]. Washington, D.C.: The Bureau [producer and distributor].

U.S. Bureau of the Census. (1993) *County Business Patterns, 1990* [computer file]. Ann Arbor, MI: ICPSR [distributor].

U.S. Bureau of the Census. (1999) *USA Counties 1998 on CD-ROM* [machine readable data file]. Washington, D.C.: The Bureau.