Spatial Wage Estimates for 19th Century US Workers*

Jason Dunn

Department of Economics, Boston University

November 25, 2026

Abstract

Measuring socioeconomic status (SES) throughout US history is challenging due to the (i) lack of direct income data and (ii) limitations of occupation scores, which ignore geographic and demographic variation. While recent literature has improved upon these scores by accounting for education, wealth, and other SES factors, they remain coarse or are limited by incomplete reporting thresholds. This paper's core contribution is developing a method for tightly estimating individual manufacturing wages from 1850 to 1880. I supplement these with broader wages on non-manufacturing occupations for greater coverage of the 19th century workforce. I then set up a procedure for benchmarking these wage estimates and other SES measures against external SES sources. I find the wage data and a composite index of wages, literacy and wealth perform at or above pre-existing SES measures. This new SES measure implies lower levels of economic mobility and greater increases in inequality over the course of the 19th century.

*Correspondence address: jtdunn@bu.edu. I am grateful to my advisor James Feigenbaum for guidance and support through this project. I also thank participants at the 2026 Southern Economic Association Conference for helpful comments and suggestions. First version: November 2026

1 Introduction

The evolution of wealth and economic mobility is a central question in economic history. Despite its importance, the precision of data for this topic is sparse for various times and locations. The direct measurement of individual income is inconsistently and non-comprehensively recorded in the past. Prevailing standard measurements of historical socioeconomic status (SES) in the US is an occupational score, which assigns a person the SES ranking of their respective occupation. This rank correlates with income and wealth, and covers any individual in the labor force.

Occupation scores are lacking along various dimensions as a proxy for SES. For one, variation in occupation score occurs only through differences in occupation, and therefore misses economic variation along demography (age, race, gender, etc.) and geography. Additionally, occupation scores do not account for changes in the relative SES of occupations over time.[1] Recent literature improves upon the standard occupation score in relation to the above issues. This includes accounting for the literacy share within race-occupation-region cells, using observable predictors such as state, sex, age, and race, or even using occupation-state average wealth values (Song et al. 2020; Ward 2023; Saavedra and Twinam 2020; Collins and Zimran 2023). While improving upon the granularity of SES measurement, these scores still lack much of the variation in income/wealth within occupation, and may not accurately proxy for SES across the full 1850–1940 time span.

In this paper, I improve nineteenth-century SES measurement by combining localized wages with literacy and wealth. The core contribution is a tightly identified manufacturing wage. For each Census of Manufactures decade from 1850 through 1880 (Hornbeck et al. 2024), I geocode establishments against that year's Census Place Project (CPP) places (Berkes, Karger, and Nencka 2023) and assign Census of Population (COP) workers an employment-weighted firm wage inside a 15-mile radius in the same industry, demographic group, and skill type. While tightly identified, manufacturing makes up only about 15% of the workforce in 1880. To speak to more general findings I therefore *supplement* the manufacturing wage with Global Price and Income History (GPIH) occupational quotes, which cover most workers (farm labor, domestics, clerks, crafts, and so on) but are typically only at the state-, region-, or national-level. I then benchmark/validate these wages against external sources at the city/county-aggregated level, and they track closely with those estimates.

Next, I construct a composite SES index by combining the spatial wage with individual

1. The relative rankings of these occupation scores were determined in 1950. The IPUMS occscore, for instance, is the median 1950 income of the occupation in hundreds of dollars.

literacy rates and local average wealth. I optimally weight the three components to maximize correlation with incomes from the Cost of Living Survey (COLS), and then compare the fit to earnings in the Historical Labor Statistics Project (HLSP), 1915 Iowa Census, and 1940 Full Count Census. With the exception of the LIDO score in 1940, the composite SES proves to track at or better than all other SES measures as correlatives to income from these sources.

Finally, I gauge the implications of this composite SES index on economic mobility and inequality. First replicating the mobility estimates of Ward 2023, I find the composite index generates higher degrees of historical persistence and lower degrees of mobility, with an improvement in mobility remaining but later into the early 20th century. Calculating Gini coefficients and $90^{th}/10^{th}$ percentile ratios, I also find that the composite SES suggests greater increases in inequality from 1850–1880 than other SES measures.

This work contributes first and foremost to an existing literature on the measurement of historical SES. Improving upon the standard ranking according to occupation score, (Song et al. 2020) construct "song scores" by calculating the share of individuals within each occupation that are marked as being literate in the census of population. They use this literacy share to proxy for the SES associated with each occupation. As a direct improvement on this measure, (Ward 2023) constructs "updated Song scores" as the literacy share by race-occupation-region cells, allowing for even further variation in these scores. While addressing additional variation masked by occupation scores, this is still a rather coarse measure of SES. To address the limitations of occupations as the only data point used for SES, (Saavedra and Twinam 2020) use an LLM on the 1950 1% sample to gauge the strongest determinants of income. From this analysis, they construct a "LIDO" score, which uses occupation, industry, and demographic information (state, sex, age, race) to estimate an SES ranking for each working individual. This improves the measure of occupation scores considerably, but the relevance of this exercise likely fades going further back in the past to 1850, when the structure of the economy was considerably different. (Collins and Zimran 2023) construct a "wealth score" which is more tailored to 19th century analysis. They calculate the average total wealth (real estate and personal property) reported in the 1870 full count decennial census of population data across state-occupation cells, and rank individuals along the percentiles of this constructed variable. This helps to address the issue of limited individual economic information in the 19th century, but wealth was only reported for households with more than $100 in total wealth, so this remains a somewhat limited proxy, particularly for the bottom end of the distribution. My constructed spatial wage estimates contribute to this growing literature on historical SES measurement by adding yet another dimension from

which to evaluate individuals' economic situations, and allowing for a finer measurement of income and wealth proxies through wages alongside localized literacy and wealth estimates.

This study also relates to an extensive literature measuring US wages in the 19th century, as occupational/establishment averages (Lebergott 1964; Weeks 1886), in Army civilian payrolls (Margo and Villaflor 1987; Margo 2000), for women and children (Goldin and Sokoloff 1982), or across regional labor markets (Sundstrom and Rosenbloom 1993; Rosenbloom 2002). (Atack, Bateman, and Margo 2004) look at a sample of the COM manuscripts to gauge the degree of wage dispersion from 1850–1880. In total, these studies measure wages from either an aggregate, regional, or firm sample. The full COM firm-level manuscripts digitized by Hornbeck et al. (2024) make the present study possible, in assigning individual-level wages and gaining a tighter understanding of wage inequality and persistence.

This paper also contributes to a directly related literature on intergenerational economic mobility (Feigenbaum 2018; Salisbury 2014; Zimran 2024; Collins and Wanamaker 2022). From an aggregate perspective, (Song et al. 2020) was the prevailing paper on the evolution of US mobility from 1850–1940, finding that mobility has decreased over time. (Ward 2023) developed a substantially important adjustment to the measurement by (Song et al. 2020), cleaning for measurement error in the father's occupation by instrumenting with a second occupation observation, and also including black individuals in the estimates. These adjustments reversed the story of mobility in the US, into a trend of increasing mobility over time. The findings of this paper can further contribute to this discussion by evaluating how the mobility of workers evolved over the course of the 19th century. Relatedly, this paper's findings also contribute to studies on inequality in the US (Soltow 1969; Steckel and Moehling 2001; Lindert and Williamson 2016; Piketty and Saez 2003). Due to the limited collection of individual income data in the 19th century, my composite SES score provides a valuable snapshot of inequality in the US during the late 1800's.

The rest of the paper is organized as follows: Section 2 outlines the various datasets used in the paper. Section 3 describes the panel spatial-wage construction for 1850–1880, the non-manufacturing wage supplements, and the construction of the composite SES index. Section 4 validates that 1880 wage and the composite SES index against contemporary earnings. Section 5 explores the implications for intergenerational mobility and inequality. Section 6 concludes.

2 Data

The data partitions into two main groups. The first group builds the panel wage and the composite SES score, while the second group validates and scores the constructed 1880 wages and composite index. All sources are described below; the panel wage construction and the composite index themselves are in Section 3.

2.1 Spatial Wage Construction

2.1.1 Census of Manufactures

The Census of Manufactures (COM) covers establishments producing at least $500 in annual value of goods in 1850, 1860, 1870, and 1880.[2] Hornbeck et al. (2024) located, digitized, and transcribed the manuscripts on microfilm.[3] Each establishment has a state, county, township, nearest post office, industry, employment level, and wages.[4]

Table A1 of the Appendix describes the 1880 file. Employment is reported for 202,419 establishments. Female employment is reported at 20,696 establishments and child employment at 16,335. The average daily wage is $1.85 for skilled workers and $1.05 for unskilled workers. Figures A3 and A4 of the Appendix map the county pattern, and Figure A1 of the Appendix plots wages by region and industry. Skilled wages sit above unskilled wages inside each region and inside each industry. Nominal wages are highest in the West and the Midwest, the regions of largest growth during this period. Construction, machinery, and printing are the highest-paying industries.

I am able to geolocate the majority of establishments by matching its township or post office to that year's historical post offices (Blevins 2021), Census Place Project towns (Berkes, Karger, and Nencka 2023), and the Geographic Names Information System (GNIS).[5] With the firms located spatially, I can link them to nearby individuals.

2.1.2 Census Place Project

The Census Place Project (CPP) assigns a consistent subcounty place to people and households in the full-count censuses (Berkes, Karger, and Nencka 2023). Places are built from

2. Some state-year manuscripts are missing. Coverage is documented at https://cmfdata.org/about.
3. The transcription is described at https://cmfdata.org/methodology.
4. The exact demographic coverage varies by year. Wages are available by gender for 1850 and 1860, as an establishment average in 1870, and as skilled and unskilled daily rates in 1880. I make use of these relative ratios to impute wages by sex and skill in each year (see more discussion of this in Section 3).
5. In 1880, 183,633 of 208,309 establishments (88.2 percent).

the National Historical Geographic Information System (NHGIS) and the GNIS and they are clustered so that a new town line or an annexation does not split one community into two places.[6] The release includes the coordinates of each place. Where the census records a place name, the match to a town or city is complete. People with no place name, often on isolated farms, are not given a subcounty point.

2.1.3 Census of Population

The population censuses used are the full counts for 1850, 1860, 1870, 1880, and 1940 (Ruggles et al. 2024). The nineteenth-century files contribute occupation, industry, sex, age, race, literacy, and, in 1850, 1860, and 1870, real estate and personal property. The 1940 file contributes the same demographic and occupation variables as well as wage-and-salary income for every worker. That income is nominal and topcoded at $5,000.

Table A2 of the Appendix describes the 16.7 million members of the workforce in the 1880 summary file. Mean age is 35. Eighty-four percent are men, 85 percent are white, 40 percent live on farms, and 32 percent live in urban places. The IPUMS occupation scores on that file, described below, are on very different scales.

2.1.4 GPIH wage quotes

Manufacturing is the well-identified end of the wage spectrum in this paper, but it is a small subset of the labor force. For workers with no COM firm record I supplement with Global Price and Income History (GPIH) quotes—the Lindert-Williamson worksheets for 1850, 1860, and 1870. That coverage comes at a cost: GPIH quotes are typically state, region, or national, not local. A quote is an occupation or an industry, a geography, a cash wage, the time unit of that wage, and, when the sheet records it, a board allowance. The extract contains 1,483 quotes. Occupations that line up with the census include farm labor, common labor, domestics, clerks, clergy, teamsters, lawyers and judges, teachers, professors, civil engineers, stationary engineers, seamen, the skilled crafts, miners, lumbermen, physicians, and librarians.

6. Places i and j are combined when they lie within $100 \times K_{cluster} \times \max\{sharepop_i, sharepop_j\}$ miles, where $sharepop_i$ is place i's share of the mapped population across years and $K_{cluster} = 5$. See Figures 4–6 in Berkes, Karger, and Nencka (2023).

2.2 Validation/Comparison Data

The sources below benchmark the constructed 1880 wage and the composite index built off of it. I do not re-run COLS, DOL, or HLSP validation for 1850–1870. Those earlier panel wages enter the 1850–1880 inequality series, and decade-varying mobility.

2.2.1 Cost of Living Survey

The BLS Cost of Living Survey (COLS) reports incomes for urban households working in manufacturing in 1888–1890. Table A3 of the Appendix covers 6,627 households. Fifty-four percent of the heads are native-born, families average 4.8 members, the husband's mean income is $589, and total household income averages $761. The file also records the spouse's income, ages, state, and the head's industry. The comparison in this paper uses the head's income, matched to the published 1880 panel wage. The cell-level income-proxy sample is the 4,121 households that fall in 170 state × demographic group × skill × industry cells with a wage, a literacy share, and wealth. Figure A6 of the Appendix shows the same broad geography and industry pattern as the 1880 manufacturing wages: higher incomes in the West and the Midwest, and higher incomes in machinery.

2.2.2 Department of Labor city wages

The Department of Labor published average wages by city and occupation from 1870 through 1898, for named crafts and laborers: bricklayers, carpenters, machinists, blacksmiths, compositors, railroad occupations, and others. The file has 8,108 city-occupation-year cells. The mean wage is $2.53. The wage-validation sample pools 1870 and 1880 into 162 city × skill × industry cells. It does not cover workers outside those cities or occupations. Appendix Figure A7 maps the cities.

2.2.3 Historical Labor Statistics Project

The Historical Labor Statistics Project (HLSP) assembles the surviving surveys of individual workers taken by state bureaus of labor statistics (Carter, Ransom, and Sutch 1991). I use 28 of those microdata files, 64,855 workers, surveyed from 1879 through 1903 (Table A4 of the Appendix). A row is a surveyed worker. The surveys are concentrated in manufacturing, with additional files for teachers, railroad workers, construction, farm labor, and domestic service. Geography is the finest place the survey coded: a town, otherwise a county, otherwise the state. The variables used here are the worker's own wage and own annual earnings,

occupation, industry, sex, and age. Firm canvasses, published occupation averages, farm-operator schedules, and owner-drivers are not worker microdata and are left out of the worker file. Dollar amounts are stored as published, which I convert to 1880 dollars with the Coelho-Shepherd index Coelho and Shepherd 1974. The daily-wage comparison uses 12,236 manufacturing workers in 375 county × demographic × skill × industry cells and 24,483 workers in 910 cells once non-manufacturing industries with a GPIH quote are added. The earnings comparison uses 14,625 manufacturing workers in 546 cells and 24,721 workers in 919 cells.

2.2.4 Iowa 1915 and the 1940 census

Two later earnings files are also considered. The 1915 Iowa state census, digitized by Goldin and Katz, contributes about 16,000 earners with an occupation that can be placed in an 1880 county × demographic cell (3,730 in manufacturing titles, 15,987 in the all-worker cell). The 1940 full-count wage-and-salary file contributes the earners below the $5,000 topcode. The rank-slope sample is 3,815,030 workers in 81,106 Census Place Project place × demographic group × industry cells.

2.2.5 Other SES scores

The population-census files carry the IPUMS occupation scores: the occupational income score, the Duncan socioeconomic index, Siegel's prestige score, occupational earnings, occupational education, and the Nam-Powers-Boyd score.[7] Four scores from the literature are used as well:

- **Song score (Song et al. 2020).** The percentile rank of an occupation's literacy share, by micro-occupation and birth cohort.
- **Adjusted Song score (Ward 2023).** The same literacy ranking inside race, occupation, and census region.
- **LIDO (Saavedra and Twinam 2020).** A score fit from occupation, industry, state, sex, age, and race, using 1950 census 1% sample.
- **Collins-Zimran wealth score (Collins and Zimran 2023).** Mean total property in 1870 by occupation and state. Cells with fewer than 100 men are filled with the broad occupation group.

7. For more information, see https://usa.ipums.org/usa/chapter4/sei_note.shtml.

2.3 Miscellaneous Data

2.3.1 Price Data

To make wages across time and place comparable, I use the regional price index of Coelho and Shepherd (1974), referenced in Table 2 of their paper. The index runs from 1851 to 1880, and covers 6 regions, in addition to a national index: New England, the Middle Atlantic, the East North Central, the West North Central, the South Atlantic, and the East South Central. I use these price indices to express the GPIH wage quotes in 1880 terms, and make all wages comparable across regions.

2.3.2 Linked fathers and sons

The main mobility sample is the father-son file prepared for Ward (2023). A link records the son's birth cohort, the occupations of the son and of the father, a second observation of the father's occupation, race, age, census region, and the sample weight. The file also carries the adjusted Song ranks used in that paper. Sons are born between 1840 and 1910.

3 Wage construction

In this section, I describe the panel spatial wages that enter the SES score, and how the 1880 cross-section of that panel lines up with higher-aggregated wages.

Figure 1 displays the intuition. Blacksmiths in the 1880 population census are plotted at their CPP location in Massachusetts, and blacksmithing firms are triangles, darker where the firm pays more. A blacksmith in Western Massachusetts should draw a lower wage than one near Boston, because the map places them nearer lower-paying blacksmithing firms. That within-occupation geographic variation is what the spatial wage is designed to recover.

3.1 Panel firm wages and the 15-mile cell

For each COM decade I first build firm-level demographic×skill wages in 1880 regional dollars. Establishments are geocoded to subcounty points by matching township and post-office names to that year's historical post offices, CPP places, and GNIS locations. I then compute a place-to-firm distance matrix for that year.

For a manufacturing worker i at location l, in industry j, demographic group g, and skill s, the baseline wage is the employment-weighted mean of firm daily wages inside radius d:

$$\omega_{i,l,j,g,s} = \frac{\sum_{f \in F_{l,j,d}} (W_{f,j,g,s} \cdot L_{f,j,g})}{\sum_{f \in F_{l,j,d}} L_{f,j,g}}. \tag{1}$$

$F_{l,j,d}$ is the set of firms in industry j within $d = 15$ miles of location l. $L_{f,j,g}$ is firm f's employment of demographic group g. The 1880 manuscript reports skilled and unskilled daily wages, and employment of men, women, and children, as separate margins: skill is not crossed with sex or age. The weight in (1) is therefore the demographic headcount, and $W_{f,j,g,s}$ is the firm's wage for skill s. A firm with no workers in the group, or with no wage for the skill, is left out of the average. The index g selects the headcount and the firms that enter the average. To partition wage by sex in 1870 and 1880, I tilt the establishment's skilled or unskilled wage by the female/male wage ratio ρ from 1850/60: $W_{f,j,\text{female},s} = \rho_{j,r} W_{f,j,s}$. Female employment is reported at 20,696 of the 202,419 establishments, and only those firms enter the female average, weighted by female employment. I classify a worker under 18, of either sex, as a child. Child employment is reported at 16,335 establishments. Children are weighted by that reported employment and assigned the unskilled rate.

Coverage of skill type and gender varies across COM census year.[8] To address this, I impute relative values in the years they are not available. Let σ_j be the skilled share of manufacturing workers in industry j in that year's population census, $\pi_{j,r}$ the 1880 skilled/unskilled ratio, and $\rho_{j,r}$ the 1850/60 female/male ratio. In 1850 and 1860 the manuscript gives sex-specific annual wages w_g; the skill split that adds up to w_g is $w_u = w_g/(\sigma_j \pi_{j,r} + 1 - \sigma_j)$ and $w_s = \pi_{j,r} w_u$. In 1870 only the establishment average $\bar{w}$ is observed; I first recover sex-specific means $w_m = \bar{w}/((1 - s_f) + s_f \rho_{j,r})$ and $w_f = \rho_{j,r} w_m$ from the firm's female employment share s_f, then apply the same skill split. The spatial average in (1) then proceeds as in 1880, with employment weights $L_{f,j,g}$ from the manuscript.

Industry is the COM title list. Exact COP, COLS, DOL, and HLSP titles are concorded onto that list (Appendix Tables A5, A6, and A7). Skill type follows a fixed occupation mapping—named professional and craft titles are skilled; named unskilled titles and residual occupations are unskilled—assigned from the COP occupation (Table A8). In 1880 only, residual census industries that do not match a COM title exactly may take a local residual match. These matches are restricted to a list of unique leftover COP titles paired to leftover COM product synonyms. Matching is done on the same priority as wages: residual COM employment within 15 miles, then county leftovers. Exact title matches are never overwritten.

Firm annual wages are winsorized at the 1st and 99th percentiles and put into 1880

8. Specifically, male and female wages are reported in 1850 and 1860, while skilled and unskilled wages are reported in 1880.

regional dollars with the (Coelho and Shepherd 1974) indices before the demographic×skill split. Robustness variants also test dividing large manuscript dailies by 100 and alternative winsor rules. When workers are later averaged to a cell, the weight is the inverse variance of the nearby firm wage, with a ridge of \$0.10 so a constant local wage still has finite weight:

$$w = \frac{1}{\mathrm{sd}(\omega)^2 + 0.1^2}. \tag{2}$$

3.2 Assignment hierarchy

A COP worker is joined to the panel wage at place × industry × demographic group × skill. When that exact cell is missing, assignment follows a fixed ladder on the same panel geocode graph:

1. Exact 15-mile place × industry × demographic × skill cell from (1).
2. If the industry is present within 15 miles but the worker's demographic×skill cell is empty, fill from another demographic×skill cell at that place × industry, preferring adult male over adult female over child, and unskilled over skilled.
3. If no firm in the worker's industry lies within 15 miles, use the county × industry × demographic × skill mean.
4. If the county cell is also missing, use the state × industry × demographic × skill mean.
5. If no manufacturing wage remains, the worker later receives a GPIH occupational quote.

County and state fills multiply the precision weight in (2) by one quarter, so a fallback wage counts less than a 15-mile wage with the same dispersion. A worker whose industry is nearby but who lacks any demographic×skill fill is left missing on the manufacturing wage; I do not substitute another industry's rate.

That wage is still a cell mean. Two manufacturing workers who share a Census Place Project place, industry, demographic group, and skill receive the same spatial wage, even when their own establishments would have paid them differently. The construction recovers within-occupation variation by moving from a national occupation score to local geography and by allowing the rate to differ by industry, demographic group, and skill. What it does not recover is inequality inside that cell: pay gaps across firms or workers within the same place–industry–demographic–skill group are averaged away in (1). Inequality and mobility results

that use the manufacturing wage should be read as variation across those cells. Because within-cell wage dispersion is omitted, inequality measures based on the spatial wage—and on the composite that uses it—are lower-bound estimates of the inequality that would be observed if every manufacturing worker carried an individual earnings rate.

3.3 Supplementing outside manufacturing

The manufacturing spatial wage is the paper's main methodological contribution. Results that stay inside manufacturing—wage validation against DOL and HLSP, manufacturing income-proxy slopes, and manufacturing Ginis—rest on that local firm construction alone. Results for the full workforce sit further along the coverage spectrum: a worker with no manufacturing wage receives a GPIH occupational quote. I treat that step explicitly as a supplement for generalized findings, not as an equally tight spatial estimate. The quote is converted to a daily cash rate for comparability to the manufacturing wage.[9] The daily rate is then put in 1880 regional dollars with the (Coelho and Shepherd 1974) index. A daily rate below 5 cents or above 20 dollars is dropped. Because GPIH has no 1880 worksheet, an 1880 worker takes an 1870 quote when one exists, then 1860, then 1850. At each quote year the geography order is the state, the Coelho region, and the United States. The same schedule fills non-manufacturing occupations in the all-worker composite and in the Ward covered-wage series. Where geography in the quote is coarse, within-occupation local variation remains unidentified—the tradeoff for covering the majority of workers.

Table A9 of the Appendix records who receives which wage in the 1880 labor force of 16,671,732 workers. A 15-mile manufacturing wage covers 1,271,415 workers (7.63 percent). County, state, and other manufacturing fallbacks add 641,938, 536,732, and 124,199 workers. GPIH quotes cover the rest of the assigned sample: 10,580,489 at the state (63.46 percent), 231,522 at the region, and 170,183 at the national level. Of the state row, 9,986,633 workers have a reported quote and 593,856 are rank-filled. The remaining 3,115,254 workers (18.69 percent) are uncovered. Table A10 of the Appendix lists the occupation titles inside the manufacturing and GPIH rows of the composite.

9. More specifically, a weekly wage is divided by 6, a monthly wage by 26, and an annual wage by 300. An hourly wage is multiplied by the hours on the sheet, or by 10 hours when the sheet does not say.

3.4 The 1940 real wage

To compare the published 1880 wage to 1940 pay, I use wage-and-salary income from the full-count census. That income is nominal and topcoded at $5,000. Before any price adjustment I drop nonpositive incomes, the missing and not-in-universe codes, and every earner with a recorded income of at least $5,000. Dropping the topcode first keeps the real censoring point from depending on the local price index. It also removes the upper tail before any score is scored. Occscore is the percentile of an occupation's median income, and LIDO is a predicted-earnings score in occupation, industry, and demographics, so both spread mass into that tail. I do not refit either score on the earners who remain below $5,000. The cell still carries the average of the score as constructed on the full earnings distribution, and that average is correlated with a 1940 mean taken only from earners below the topcode. Each remaining income is deflated by the 1935 WPA index of the worker's county, which corresponds to the closest city in the price indices.[10]

Manufacturing cells are the Census Place Project place, the demographic group, and the industry. Skill is averaged away, with the precision weights from equation (2), so the published 1880 wage and the 1940 income meet on the same grain. The all-worker file uses the same price adjustment and, where the 1880 worker has no manufacturing wage, fills the 1880 side with the GPIH quote.

3.5 Wage Comparison Results

I validate the 1880 panel wage against Department of Labor city wages and against HLSP daily wages. Each variable is replaced by its midrank at the cell. DOL pools 1870 and 1880 city × skill × industry cells and is unweighted. HLSP is a county × demographic group × skill × industry cell, and the rank weights the cell by the number of workers with a daily wage. The manufacturing HLSP sample keeps mapped COM industries. The all-worker sample adds agriculture, domestic service, hack and bus, teachers, clerks, and laborers, named to a GPIH quote.

Figure 2 plots binscatters of those percentiles. The relationship is positive in every panel. The rank correlation is 0.61 with the Department of Labor wage (162 cells), 0.76 with the HLSP manufacturing daily wage (375 cells), and 0.85 with the HLSP daily wage for all scored workers (910 cells).

10. More specifically, the county is given the nearest of the 59 cities, using the county internal point. When a second city is within five miles and the nearest city is not at the county point itself, the index is the inverse-distance average of the two, $(P_1/d_1 + P_2/d_2)/(1/d_1 + 1/d_2)$.

Figure 1: Blacksmith Residences and Firm Locations

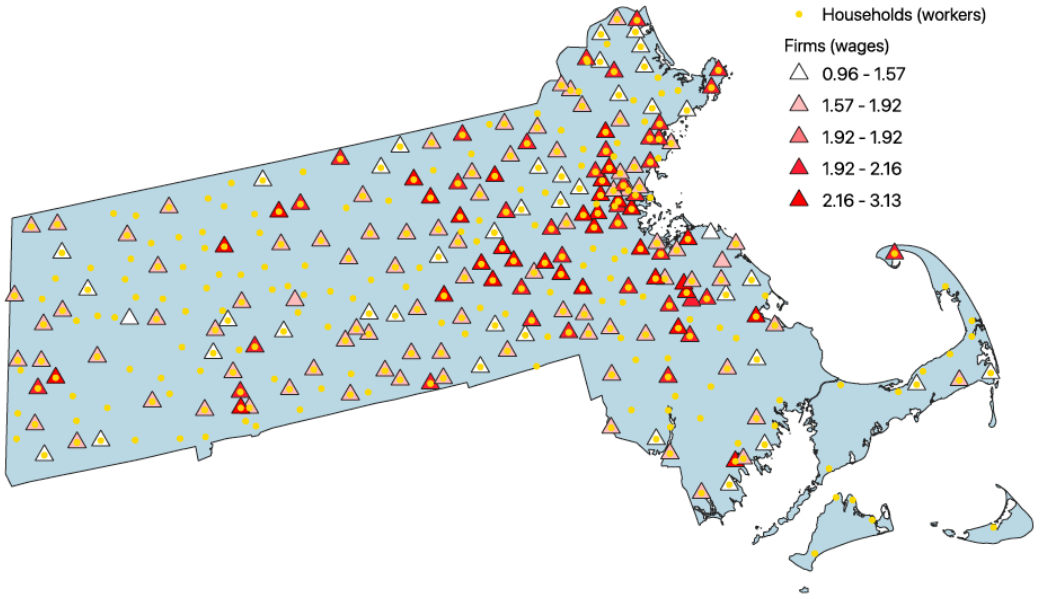


Figure 2: Spatial wage against external pay

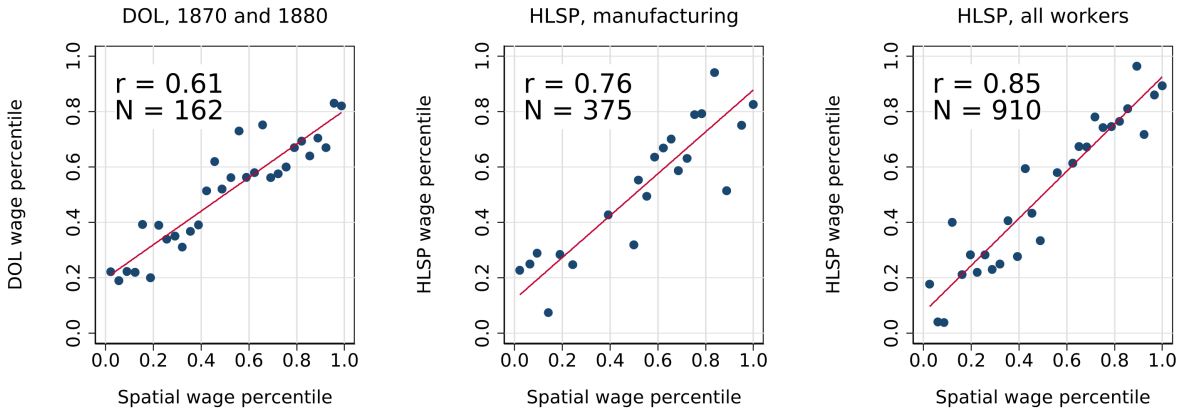


Table 1 puts the same three outcomes against the other SES scores. The DOL column is the rank correlation. The HLSP columns are rank slopes of the cell-mean daily wage on the score. The spatial wage is the strongest correlate of the DOL wage (0.61). Occscore and occupational earnings are 0.27, LIDO is 0.20, and the literacy and prestige scores are near zero or negative. Inside HLSP manufacturing the spatial-wage slope is 0.76, tied with LIDO, Siegel prestige, and occupational education, and above occscore (0.69). On all HLSP workers the spatial-wage slope is 0.85, next to LIDO (0.84) and above occscore (0.80). Collins-Zimran wealth is essentially uncorrelated with the DOL wage (0.01) and negatively related to HLSP pay (-0.66 in manufacturing). Song, adjusted Song, Collins-Zimran wealth, and LIDO are observed on the 1880 DOL cells only (90 cells). The other occupation scores are on 148 DOL cells. LIDO is missing for 32 manufacturing cells and 95 all-worker cells.

Table 1: External wages and SES scores

	DOL, 1870–1880	HLSP manufacturing	HLSP all workers
Spatial wage	0.61	0.76	0.85
Occscore	0.27	0.69	0.80
Occupational earnings	0.27	0.69	0.81
LIDO	0.20	0.76	0.84
Nam-Powers-Boyd	0.06	0.06	0.64
Adjusted Song	0.04	0.54	0.39
Collins-Zimran wealth	0.01	-0.66	-0.25
Siegel prestige	-0.10	0.76	0.36
Occupational education	-0.12	0.76	0.45
Duncan SEI	-0.15	0.43	0.59
Song	-0.20	0.61	0.44
Cells	162	375	910
Workers	–	12,236	24,483

The DOL column is the rank correlation. The HLSP columns are rank slopes of the cell-mean daily wage on the score. Rows are ordered by the DOL correlation. DOL is one city × skill × industry cell in 1870 or 1880, ranked across both years. HLSP is a county × demographic group × skill × industry cell, ranked with weights equal to the number of workers. The spatial wage in the DOL column is in 1880 dollars. Song, adjusted Song, Collins-Zimran wealth, and LIDO are on the 1880 DOL cells only (90 cells). The other occupation scores are on 148 DOL cells. LIDO is missing for 32 manufacturing cells and 95 all-worker cells.

To gauge the sensitivity of results, Table A11 of the Appendix repeats this wage validation for various unilateral deviations from the wage construction pipeline. The results appear to be largely robust to alternative forms of weighting, alternative distance thresholds, shutting down each cell dimension, alternate wage cleaning, and bayes shrinkage.

This exercise validates the 1880 panel wage against higher-aggregated pay from the pe-

riod. I next fold that wage into a composite SES score and compare it to existing scores.

4 Validation

I now construct a composite SES index on the published 1880 panel wage, combining it with literacy and wealth to proxy nineteenth-century individual incomes. Weight choice and the main income-proxy slopes are likewise anchored on 1880 inputs; Iowa 1915 and 1940 enter only as out-of-sample robustness.

4.1 Literacy, wealth, and the composite

The composite has three inputs. The wage is the panel spatial wage. Literacy is the reading and writing ability of each individual, as recorded in the census. The wealth input is average real estate and personal property at the at the CPP place × occupation × demographic group (white men, non-white men, women) × year.[11] Wealth can be noisy at the local level, so it is shrunk toward the person-weighted mean of the census region-occupation-group.

$$y_j^{\mathrm{EB}} = \mu + B_j(y_j - \mu), \qquad B_j = \frac{\tau^2}{\tau^2 + \sigma^2/n_j}. \tag{3}$$

Where μ is the person-weighted mean in the occupation × census region × demographic group, n_j is the number of people in place cell j, σ^2 is the variance of one person, and τ^2 is the between-place variance of the cell means. Let $N = \sum_j n_j$ and G be the number of place cells in the group. Then $\mu = \sum_j n_j y_j / N$, $Q = \sum_j n_j (y_j - \mu)^2$, and

$$\tau^2 = \max\left\{0,\ \frac{Q - (G-1)\sigma^2}{N - \sum_j n_j^2 / N}\right\},$$

with $\tau^2 = 0$ when $G < 2$. $B_j = 1$ keeps the local value, while $B_j = 0$ uses the group mean. For wealth, σ^2 is the person-level variance of census property in the same occupation-region-group. Wealth uses the shrunk place mean when the place is observed in 1870, and the occupation-region-group mean otherwise. The wage cell is still child, adult man, or adult woman. Literacy and wealth follow sex and race, and they are averaged onto that wage cell with the same precision weights as the spatial wage.

11. Wealth information was not recorded in the 1880 full count census, so I use 1870 as the nearest available data for this.

Let z_w, z_ℓ, and z_a be the wage, the literacy, and average wealth, each standardized with the household-weighted mean and standard deviation on the COLS estimation sample. The composite SES index is then

$$C = \lambda_w z_w + \lambda_\ell z_\ell + \lambda_a z_a, \tag{4}$$

with $\lambda_w, \lambda_\ell, \lambda_a \geq 0$ and $\lambda_w + \lambda_\ell + \lambda_a = 1$. The weights maximize the household-weighted rank correlation between C and the household head's income in the Cost of Living Survey, so that source is used to choose the weights. The fit puts 0.51 on the spatial wage, 0.40 on literacy, and 0.09 on wealth. The in-sample rank correlation is 0.56. The same weights are used for every composite score below. I also report an equal-weighted composite that puts one third on each standardized input, and I report each of the three inputs separately.

Table 2 decomposes each score on 1880 workers who have a coded occupation, a Census Place Project place, a demographic group, and a wage. The five shares are nested and sum to one: occupation, then census region, state, and place, then demographic group and the remaining within-cell variation. The composite puts 0.44 of its variance on occupation, 0.25 on region, 0.05 on state, 0.12 on place, and 0.14 on demographic group and skill. The spatial wage is 0.85 occupation. Literacy is the most geographic of the three inputs (0.18 occupation, 0.37 region, 0.17 place). Wealth loads on occupation (0.42) and place (0.44). Adjusted Song is 0.58 occupation and 0.36 region, with almost no state or place residual. LIDO is 0.66 occupation. Occscore, Duncan SEI, Siegel prestige, Nam-Powers-Boyd, and the occupational education and earnings scores are occupation scores: their shares sit on occupation. This showcases some of the added variation this new SES information aims to provide.

4.2 Cell rank slopes

With the weights fixed on COLS, I evaluate how well each score proxies for earnings. The outcome is the mean earnings in the cell the score is attached to. Both the earnings rank and the score rank use weights equal to the number of households or workers in the cell. Let $\widetilde{y}_c$ and $\widetilde{S}_c$ be those midranks. The reported entry is the slope

$$\widetilde{y}_c = \alpha + \beta\, \widetilde{S}_c + \varepsilon_c. \tag{5}$$

A COLS cell is a state, demographic group, skill, and industry. An HLSP cell is a county, demographic group, skill, and industry. Manufacturing keeps mapped COM industries. All

Table 2: Variance decomposition of SES scores with nested geography, 1880

	Occupation	+ Region	+ State	+ Place	+ Demo/Skill
Composite SES	0.44	0.25	0.05	0.12	0.14
Equal weights	0.40	0.24	0.04	0.19	0.13
Spatial wage	0.85	0.02	0.04	0.05	0.04
Literacy share	0.18	0.37	0.06	0.17	0.22
Average wealth	0.42	0.09	0.03	0.44	0.02
Adjusted Song	0.58	0.36	0.00	0.01	0.05
LIDO	0.66	0.15	0.07	0.02	0.10
Collins-Zimran wealth	0.73	0.13	0.14	0.00	0.00
Song	0.98	0.00	0.00	0.00	0.02
Siegel prestige	1.00	0.00	0.00	0.00	0.00
Nam-Powers-Boyd	1.00	0.00	0.00	0.00	0.00
Duncan SEI	1.00	0.00	0.00	0.00	0.00
Occscore	1.00	0.00	0.00	0.00	0.00
Occupational education	1.00	0.00	0.00	0.00	0.00
Occupational earnings	1.00	0.00	0.00	0.00	0.00

Each entry is the share of that score's variance; the five shares in a row sum to one. Columns are nested: occupation, then census region, state, and Census Place Project place, then demographic group and remaining within-cell variation (industry and skill). Pure occupation scores load on occupation. Adjusted Song's geographic content is mostly regional; the spatial wage and Composite SES load more on local place. Workers have a coded occupation, a place, a demographic group, and a wage (Census of Manufactures or a GPIH quote). The Composite SES and equal-weight index use weights 0.51, 0.40, and 0.09 on the standardized spatial wage, unshrunk literacy share, and Bayes-shrunk average wealth, and are observed for 13,554,037 workers. LIDO is observed for 9,880,595 workers.

workers adds the GPIH industries named above. The composite, the equal-weight index, and wealth require census wealth, which those GPIH industries do not have, so those three rows are estimated on the cells where wealth is observed. Iowa 1915 is a county × demographic group cell. Both 1940 columns are a Census Place Project place × demographic group × industry cell; manufacturing drops non-manufacturing industries.

Table 3: Cell rank slopes of earnings on SES scores

	COLS	HLSP		Iowa 1915		1940	
	training	mfg.	all	mfg.	all	mfg.	all
Composite SES	0.56	0.75	0.75	0.25	0.28	0.35	0.35
Equal weights	0.53	0.76	0.76	0.17	0.31	0.36	0.36
Spatial wage	0.38	0.78	0.85	0.68	0.40	0.30	0.30
Literacy share	0.42	0.12	0.42	-0.51	-0.38	0.18	0.18
Average wealth	0.29	0.46	0.46	0.14	0.29	0.30	0.30
Adjusted Song	0.51	0.54	0.42	-0.28	-0.08	0.33	0.33
Siegel prestige	0.49	0.72	0.37	-0.24	-0.13	0.26	0.26
Collins-Zimran wealth	0.47	-0.49	-0.22	-0.05	0.09	0.25	0.25
LIDO	0.44	0.79	0.85	0.17	0.15	0.62	0.62
Nam-Powers-Boyd	0.42	0.26	0.67	-0.09	-0.30	0.41	0.41
Duncan SEI	0.38	0.48	0.62	-0.20	-0.37	0.37	0.37
Occscore	0.37	0.73	0.81	-0.06	-0.13	0.38	0.38
Occupational education	0.31	0.76	0.48	-0.20	-0.28	0.27	0.27
Occupational earnings	0.31	0.73	0.81	-0.10	-0.15	0.31	0.31
Song	0.26	0.59	0.46	-0.26	-0.08	0.21	0.21
Cells	170	546	919	31	34	81,106	81,106
Households / workers	4,121	14,625	24,721	3,730	15,987	3,815,030	3,815,030

Each entry is the rank slope of cell-mean earnings on the score. Cells are ranked with weights equal to the number of households (COLS), workers (HLSP), earners (Iowa 1915), or wage earners (1940). Literacy is the unshrunk share of individuals who read and write. The Composite SES refits the COLS weights on that share: 0.51, 0.40, and 0.09 (COLS household-weighted rank correlation 0.56). COLS is a state × demographic group × skill × industry cell. Both HLSP columns are a county × demographic group × skill × industry cell. Manufacturing keeps mapped COM industries. All workers adds agriculture, domestic service, hack and bus, teachers, clerks, and laborers, named to the GPIH quote (farm labor, female domestics, teamster, teachers, clerks, common labor), whose wage is that quote. The composite, equal-weight, and wealth scores require census wealth, which the GPIH industries do not have. Iowa 1915 earnings are from the Goldin–Katz digitization of the Iowa state census. Both Iowa columns are a county × demographic group cell; manufacturing keeps production occupation titles. Both 1940 columns are a place × demographic group × industry cell; manufacturing drops non-manufacturing industries.

The composite index outperforms all other SES indicators in the training set, as well as most indicators in the HLSP sample. On the COLS training cells the composite slope is 0.56, compared to equal weights (0.53), adjusted Song (0.51), Siegel prestige (0.49), Collins-

Zimran wealth (0.47), and LIDO (0.44). The spatial wage alone is 0.38 and the literacy share is 0.42. On HLSP manufacturing earnings the composite slope is 0.75. The spatial wage is 0.78 and LIDO is 0.79. Equal weights are 0.76, occscore is 0.73, and adjusted Song is 0.54. Literacy falls to 0.12. Collins-Zimran wealth is −0.49. On all HLSP workers the composite and equal weights are 0.75 and 0.76, the spatial wage and LIDO are 0.85, and occscore is 0.81.

Similar results tend to prevail for Iowa and the 1940 full count, though the interpretation here should be second order considering the later time period. On Iowa manufacturing earnings the spatial wage has the steepest slope (0.68). The composite is 0.25 and literacy is −0.51. On all Iowa earners the composite is 0.28, equal weights 0.31, and the spatial wage 0.40. Literacy and most occupation scores are negatively related to Iowa earnings. Against 1940 income the composite slope is 0.35 in both the manufacturing and all-worker columns (81,106 cells and 3,815,030 workers in each). LIDO is 0.62. Occscore is 0.38 and adjusted Song is 0.33. The 1880 composite still ranks 1940 pay, and LIDO, which is built from later earnings correlates, ranks it more closely.

5 Inequality and mobility

I next use the composite index to measure inequality changes in the 19th century and to revisit intergenerational mobility estimates in the linked sample of (Ward 2023).

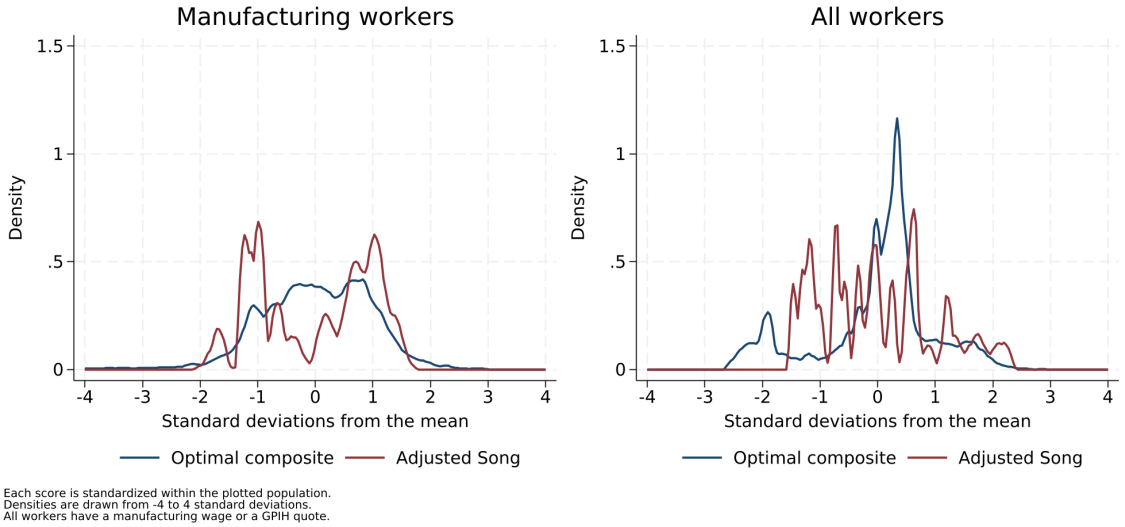
5.1 Gini coefficients

I first compare the shape of the 1880 scores. Figure 3 plots the standardized composite against the adjusted Song score. The adjusted Song score bunches. The composite is smoother, because the wage and local wealth move inside the occupation. This showcases some of the added information the composite index provides relative to the adjusted song, for one alternative SES score.

To compare with other SES scores more broadly, I put every score on a common scale by standardizing to mean zero and standard deviation one inside the sample, capping it at four standard deviations, and exponentiating.[12] Table A12 of the Appendix reports the Gini coefficients in 1880 for manufacturing and all workers, at the individual, state, and regional levels. In terms of estimates of inequality within manufacturing, the composite SES sits

12. The Gini of a normal distribution on this scale is 0.52.

Figure 3: Distributions of standardized SES scores, 1880



roughly in the middle of the set of pre-existing SES indicators. It provides an estimate on the lower end for all workers, which is consistent with the broader scale on non-manufacturing wages.

As the exact level of inequality estimates is less informative when comparing across SES measures covering different information, I instead focus more closely on the relative change in inequality projected within each of these measures. Figure 4 shows the change in inequality implied by the composite SES from 1850 to 1880, for both manufacturing and all workers, and as a gini coefficient and $90^{th}/10^{th}$ percentile ratios. All four of these series indicate an increase in inequality over this time period, with varying degrees of magnitude. Table 4 reports this change for the manufacturing workforce gini coefficients for the full set of SES measures. Comparing to other measures, the composite index implies a greater increase in inequality over this time period, and this seems to be driven particularly by wage inequality and wealth inequality. This is consistent with the finding of increasing wage dispersion in manufacturing wages over this time span from (Atack, Bateman, and Margo 2004). Table A13 of the Appendix reports the same coefficients for the whole workforce, and the results are largely the same. Table A14 of the Appendix shows similar results using the $90^{th}/10^{th}$ percentile ratios.

Figure 4: Composite inequality, 1850–1880

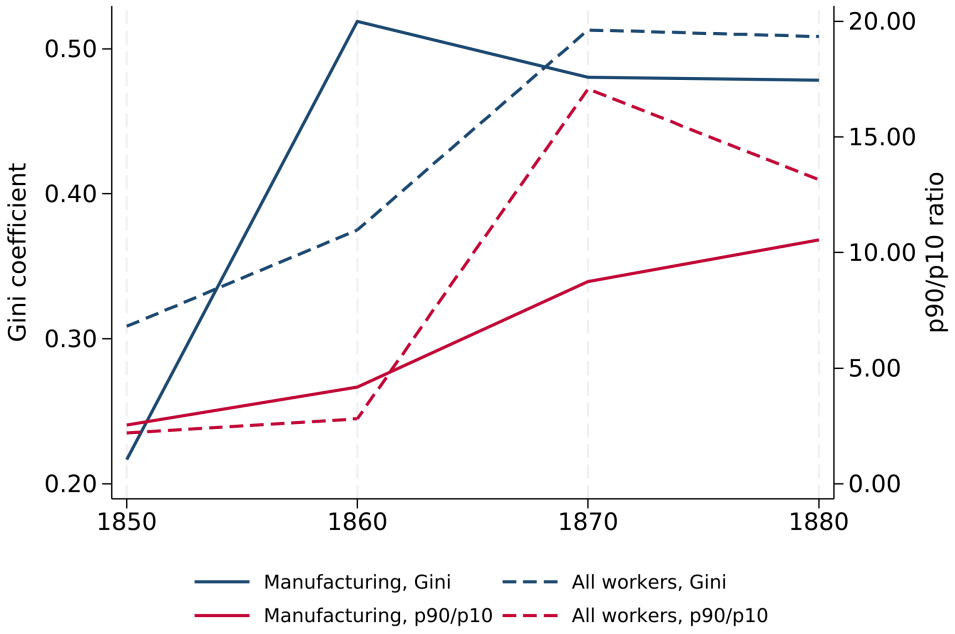


Table 4: Manufacturing Gini coefficients, 1850–1880

	1850	1860	1870	1880	1880−1850
Composite SES	0.22	0.52	0.48	0.48	0.26
Equal weights	0.21	0.56	0.53	0.57	0.36
Spatial wage	0.26	0.65	0.54	0.55	0.29
Average wealth	0.23	0.64	0.58	0.67	0.43
Nam-Powers-Boyd	0.57	0.55	0.50	0.54	-0.02
Collins-Zimran wealth	0.66	0.58	0.58	0.66	0.00
Duncan SEI	0.70	0.66	0.66	0.71	0.01
Occscore	0.55	0.52	0.52	0.59	0.05
Occupational earnings	0.44	0.45	0.45	0.51	0.07
Occupational education	0.53	0.49	0.52	0.61	0.08
LIDO	0.36	0.37	0.39	0.46	0.10
Siegel prestige	0.45	0.47	0.50	0.57	0.12
Adjusted Song	0.26	0.29	0.37	0.44	0.18
Song	0.18	0.22	0.32	0.38	0.20

5.2 Father-son persistence

Next I replicate, and reproduce with the composite SES, the intergenerational mobility estimates calculated in (Ward 2023). That paper provides a methodological improvement upon the pre-existing literature of intergenerational mobility estimates by implementing the following updates to the standard Song et al. 2020 correlations:

- Instrument for father's occupation with a second observation of that father's occupation in an adjacent census, to clean measurement error
- Measure SES (literacy shares) by race-occupation-region cells
- Include southern blacks in 1870-1880 by imputing their fathers as slaves

In order to make these improvements, Ward 2023 constructs an intergenerational linked dataset of the US from 1850–1940 with the Census Bureau's full count individual data. All sons age 0-14 are identified in the 1850–1940 full count samples and linked to their fathers in the same census. Next, sons and fathers are linked to themselves in additional censuses using off-the-shelf census links from the Census Linking Project (Abramitzky et al. 2025). For sons, links are attempted 20, 30, and 40 years forward from the son-age 0-14 census, where the observation closest to son-age 40 is used for SES outcomes. For fathers, links are attempted for the nearest census above and below the son-age 0-14 census. Mobility is then calculated as father-son rank associations from the following two-stage least squares model:

$$\text{Rank}_{i,c}^{\text{son}} = \alpha_c + \beta_c \widehat{\text{Rank}}_{f,c}^{\text{father}} + \varepsilon_{i,c} \tag{6}$$

$$\text{Rank}_{f,c}^{\text{father}} = \gamma_c + \delta_c \text{Rank}_{f,c,t+/-1}^{\text{father}} + \nu_{f,c} \tag{7}$$

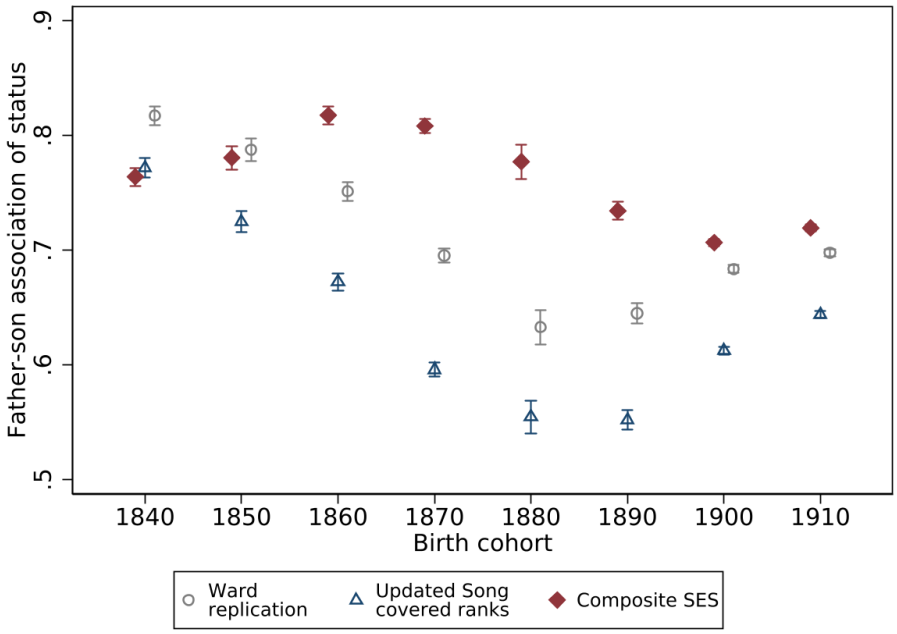
where $\text{Rank}_{i,c}^{\text{son}}$ is the adjusted percentile rank of son i in cohort c, $\widehat{\text{Rank}}_{f,c}^{\text{father}}$ is the predicted father's rank from the first stage, $\text{Rank}_{f,c,t+/-1}^{\text{father}}$ is the instrument (father's status at alternative census year), and β_c is the father-son association coefficient (the coefficient of interest).

The baseline composite for mobility is decade-varying and draws manufacturing wages from the same 1850–1880 panel described in Section 3. To construct a time series of the composite value, I use its subcomponents in the years they are available and fix to the most

recent year thereafter.[13] For robustness, I fix each component to their latest available year for the ranking through the whole series, and the results are unchanged.

In this exploration, I report three series in Figure 5. The first is the replication of the adjusted Song score on the full linked sample from (Ward 2023). The second is that same score on the links where the composite SES index is observed for the son and for both father observations. The third replaces the adjusted Song score with the composite index.

Figure 5: Father-son persistence: literacy scores and wage composites



On the full linked sample the adjusted Song score falls from 0.82 for sons born in 1840 to 0.63 for sons born in 1880, and is 0.70 for sons born in 1910. That is the increasing-mobility result of Ward (2023), with a partial reversal after 1880. Restricting the Song score to the links where the composite is available lowers the estimates across all cohorts but maintains the trend. The composite SES index on those same links is more persistent throughout the whole period, and particularly in the earlier cohorts. For the 1840 cohort the association is 0.84, against 0.77 for the covered adjusted Song score. It is 0.68 for the 1880 cohort, against 0.58, and 0.70 for the 1910 cohort, against 0.66. The trend of mobility still suggests an increase in mobility over time, though it is more delayed than in the song score evaluation, and the composite score even appears to become slightly less mobile/more persistent through

13. For the spatial wages, this is fixed past 1880, for literacy past 1930, and for wealth past 1870.

the 19th century.

Figure A8 of the Appendix evaluates various alternative specifications of this mobility result. The findings are robust to fixing the rankings to 1880, excluding shrinkage, using a county-level manufacturing wage, and excluding imputed median wages for missing observations. Figure A9 of the Appendix repeats the mobility exercise for a group of other SES indicators. The determination of the overall trend during this time period appears to be very sensitive to the SES measure used, suggesting a lot of care needs to be taken into evaluating which indicator is most appropriate for measuring mobility.

The inequality series and the father-son estimates found here are consistent. The composite records a larger rise in inequality from 1850 to 1880 than standard occupation scores, and intergenerational persistence that is higher than that of the adjusted Song ranking. Mobility is lower, and the increase in inequality is larger, than other scores imply.

6 Conclusion

This paper's core contribution is a method for tightly estimating nineteenth-century manufacturing wages. I construct an 1850–1880 panel of employment-weighted firm rates within 15 miles at place × industry × demographic group × skill. Validation, composite weights, and the main income-proxy comparisons focus on the 1880 cross-section of that panel. Manufacturing is where identification is strongest, but it is a small share of the workforce. I therefore supplement with GPIH occupational quotes when speaking to all-worker inequality and mobility, while being clear that those quotes buy coverage at the cost of geographic precision. Combining the 1880 wage with local literacy and wealth either matches or improves the SES proxy relative to existing scores.

The implications of this new composite SES index are that inequality increased by more in the 19th century than suggested by other SES measures. Additionally, intergenerational mobility is lower in the past under this measure. This somewhat modifies our understanding of inequality and mobility in the 19th century US.

There are some limitations to the above approach. The manufacturing wage remains a cell mean, so within-cell firm and worker inequality is masked and manufacturing Ginis are a lower bound. Outside manufacturing, GPIH quotes often stop at the state, so much agricultural and service pay variation stays hidden. The contribution is nevertheless a usable framework: a tightly identified manufacturing wage, optionally supplemented for broader coverage, and a composite that matches/improves upon existing historical SES measures.

References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, Santiago Perez, and Myera Rashid. 2025. *Census Linking Project.*

Atack, Jeremy, Fred Bateman, and Robert A. Margo. 2004. "Skill Intensity and Rising Wage Dispersion in Nineteenth-Century American Manufacturing." *The Journal of Economic History* 64 (1): 172–192.

Berkes, Enrico, Ezra Karger, and Peter Nencka. 2023. "The census place project: A method for geolocating unstructured place names." *Explorations in Economic History,* Methodological Advances in the Extraction and Analysis of Historical Data, 87 (January): 101477. https://doi.org/10.1016/j.eeh.2022.101477.

Blevins, Cameron. 2021. *Paper Trails: The US Post and the Making of the American West.* Google-Books-ID: g68gEAAAQBAJ. Oxford University Press, March.

Carter, Susan B, Roger L Ransom, and Richard Sutch. 1991. "The Historical Labor Statistics Project at the University of California." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 24 (2): 52–65. https://doi.org/https://doi.org/10.1080/01615440.1991.9955293.

Coelho, Philip R. P., and James F. Shepherd. 1974. "Differences in Regional Prices: The United States, 1851–1880." *The Journal of Economic History* 34, no. 3 (September): 551–591. https://doi.org/10.1017/S0022050700079808.

Collins, William J., and Marianne H. Wanamaker. 2022. "African American Intergenerational Economic Mobility since 1880." *American Economic Journal: Applied Economics* 14, no. 3 (July): 84–117. https://doi.org/10.1257/app.20170656.

Collins, William J., and Ariell Zimran. 2023. "Working Their Way Up? US Immigrants' Changing Labor Market Assimilation in the Age of Mass Migration." *American Economic Journal: Applied Economics* 15, no. 3 (July): 238–269. https://doi.org/10.1257/app.20210008.

Feigenbaum, James J. 2018. "Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940." _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12525, *The Economic Journal* 128 (612): F446–F481. https://doi.org/10.1111/ecoj.12525.

Goldin, Claudia, and Kenneth Sokoloff. 1982. "Women, Children, and Industrialization in the Early Republic: Evidence from the Manufacturing Censuses." *The Journal of Economic History* 42, no. 4 (December): 741–774. https://doi.org/10.1017/S0022050700028321.

Hornbeck, Richard, Shanon Hsu, Anders Humlum, and Martin Rotemberg. 2024. *Gaining Steam: Incumbent Lock-in and Entrant Leapfrogging.* Working Paper, April. https://doi.org/10.3386/w32384.

Lebergott, Stanley. 1964. *Manpower in economic growth; the American record since 1800. –.* New York : McGraw-Hill.

Lindert, Peter H, and Jeffrey G Williamson. 2016. "Unequal gains: American growth and inequality since 1700." *Juncture* 22 (4): 276–283. https://doi.org/10.1111/j.2050-5876.2016.00874.x.

Margo, Robert A. 2000. *Wages and Labor Markets in the United States, 1820-1860.* Chicago, UNITED STATES: University of Chicago Press.

Margo, Robert A., and Georgia C. Villaflor. 1987. "The Growth of Wages in Antebellum America: New Evidence." *The Journal of Economic History* 47, no. 4 (December): 873–895. https://doi.org/10.1017/S0022050700049858.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *The Quarterly Journal of Economics* 118, no. 1 (February): 1–41. https://doi.org/10.1162/00335530360535135.

Rosenbloom, Joshua L. 2002. *Looking for Work, Searching for Workers: American Labor Markets During Industrialization.* Google-Books-ID: jyuswV3Wf5wC. Cambridge University Press, March.

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, David Hacker, Evan Roberts, and Robert Warren. 2024. *IPUMS Ancestry Full County Data.* Minneapolis, MN. https://doi.org/https://doi.org/10.18128/D014.V4.0.

Saavedra, Martin, and Tate Twinam. 2020. "A machine learning approach to improving occupational income scores." *Explorations in Economic History* 75 (January): 101304. https://doi.org/10.1016/j.eeh.2019.101304.

Salisbury, Laura. 2014. "Selective migration, wages, and occupational mobility in nineteenth century America." *Explorations in Economic History* 53 (July): 40–63. https://doi.org/10.1016/j.eeh.2014.02.001.

Soltow, Lee C. 1969. "Evidence on Income Inequality in the United States, 1866-1965." *The Journal of Economic History* 29 (2): 279–286.

Song, Xi, Catherine G. Massey, Karen A. Rolf, Joseph P. Ferrie, Jonathan L. Rothbaum, and Yu Xie. 2020. "Long-term decline in intergenerational mobility in the United States since the 1850s." *Proceedings of the National Academy of Sciences* 117, no. 1 (January): 251–258. https://doi.org/10.1073/pnas.1905094116.

Steckel, Richard H., and Carolyn M. Moehling. 2001. "RISING INEQUALITY: TRENDS IN THE DISTRIBUTION OF WEALTH IN INDUSTRIALIZING NEW ENGLAND." *The Journal of Economic History* 61, no. 1 (March): 160–183. https://doi.org/10.1017/S0022050701025086.

Sundstrom, William A., and Joshua L. Rosenbloom. 1993. "Occupational Differences in the Dispersion of Wages and Working Hours: Labor Market Integration in the United States, 1890-1903." *Explorations in Economic History* 30, no. 4 (October): 379–408. https://doi.org/10.1006/exeh.1993.1016.

Ward, Zachary. 2023. "Intergenerational Mobility in American History: Accounting for Race and Measurement Error." *American Economic Review* 113, no. 12 (December): 3213–3248. https://doi.org/10.1257/aer.20200292.

Weeks, Joseph D. 1886. *Report on the statistics of wages in manufacturing industries: with supplementary reports on the average retail prices of necessaries of life, and on trades societies, and strikes and lockouts.* [U.S. 47th Congress, 2d sess.] House. Misc. doc ;no. 33. Washington: Govt. Print. Off.

Zimran, Ariell. 2024. "Internal Migration in the United States: Rates, Selection, and Destination Choice, 1850–1940." *The Journal of Economic History* 84, no. 3 (September): 727–766. https://doi.org/10.1017/S0022050724000251.

Appendix

Table A1: Summary statistics for Census of Manufactures establishments, 1880

	N	Mean	Median	Std. dev.	Min	Max
Most annual employees	202,419	12.22	4	55.79	0	9,000
Avg male emp (> 16)	195,280	7.71	2	33.30	0	6,000
Avg female emp (> 15)	20,696	15.97	3	63.21	0	3,000
Avg child emp	16,335	17.48	2	1,112.40	0	140,000
Avg day wages (skilled)	186,009	1.85	2	0.64	1	4
Avg day wages (unskilled)	155,202	1.05	1	0.41	0	3
Total annual wages	204,412	3,671.15	700	35,667.11	0	10238087

Table A2: Summary statistics for the 1880 labor-force sample

	N	Mean	Median	Std. dev.	Min	Max
Male	16,671,732	0.84	1.0	0.37	0.0	1.0
Urban	16,671,732	0.32	0.0	0.47	0.0	1.0
White	16,671,732	0.85	1.0	0.35	0.0	1.0
Farm	16,671,732	0.40	0.0	0.49	0.0	1.0
Age	16,671,732	34.86	31.0	14.62	16.0	127.0
Literate	16,671,732	0.82	1.0	0.38	0.0	1.0
Occupational income score	16,671,732	18.71	14.0	10.27	3.0	80.0
Duncan socioeconomic index	16,671,732	19.66	14.0	18.38	3.0	96.0
Occ prestige score, Siegel	16,671,732	31.99	32.6	12.50	0.0	81.5
Occ earnings score	16,671,732	31.96	9.9	47.25	0.8	999.9
Occ education score	16,671,732	10.15	4.6	42.36	0.0	999.9
NPB occ stat score	16,671,732	31.57	18.9	45.75	0.7	999.9

Table A5: Baseline concordance of COM titles and COP industries

COM title	ind1950	COP industry	Assignment
blacksmithing	—	Unmapped industry	Occupation contains "blacksmiths"
boots and shoes	488	Footwear, except rubber	All workers
bottles	316	Glass and glass products	Occupation contains "bottle"
bread and bakery products	416	Bakery products	All workers
carpentering	—	Unmapped industry	Occupation contains "carpenters" or "cabinetmakers"
coke	477	Misc. petroleum and coal products	All workers
cotton goods	439	Yarn, thread, and fabric	All other workers in the industry
drugs, chemicals, and medicines	467	Drugs and medicines	All workers

Continued on next page

Table A5, continued

COM title	ind1950	COP industry	Assignment
flour and grist mills	409	Grain-mill products	All workers
furniture, other	309	Furniture and fixtures	All workers
glass, n.s.	316	Glass and glass products	All other workers in the industry
hosiery and knit goods	436	Knitting mills	All workers
iron and steel, forged and wrought	336, 337	Blast furnaces, steel works, and rolling mills	Occupation contains "heaters" or "puddlers"
iron and steel, general	336, 337	Blast furnaces, steel works, and rolling mills	All other workers in these industries
iron and steel, pig	336, 337	Blast furnaces, steel works, and rolling mills	Occupation contains "furnacemen" or "smeltermen," and not "heaters" or "puddlers"
lumber, sawed	307	Sawmills, planing mills, and mill work	All workers
machinery, other	—	Unmapped industry	Occupation contains "machinists"
masonry, brick and stone	—	Unmapped industry	Occupation contains "brickmasons" or "masons"
millwrighting	—	Unmapped industry	Occupation contains "millwrights"
mining, coal	216	Coal mining	All workers
mining, iron	206	Metal mining	All workers
painting	—	Unmapped industry	Occupation contains "painters"
paints	468	Paints, varnishes, and related products	All workers
patterns and models	—	Unmapped industry	Occupation contains "pattern makers"
printing, job	459	Printing, publishing, and allied industries	All workers
silk and silk goods	439	Yarn, thread, and fabric	Occupation contains "silk"
slaughter and meat packing	406	Meat products	All workers
steam engines and boilers	—	Unmapped industry	Occupation contains "stationary engineers" or "boilermakers"
tobacco, general	429	Tobacco manufactures	All workers
unclassified	399, 499	Misc. manufacturing industries	All workers
woolen goods	439	Yarn, thread, and fabric	Occupation contains "wool" and not "silk"

COM titles are the firm's own industry string, which is the wage join key. COP workers are assigned that title from `ind1950`. Occupation is used only to split a census industry across COM titles, or, in the rows marked "Unmapped industry," when the census industry has no COM title. Later rules override earlier ones, so the assignment column is the resulting partition.

Figure A1: COM wage distribution

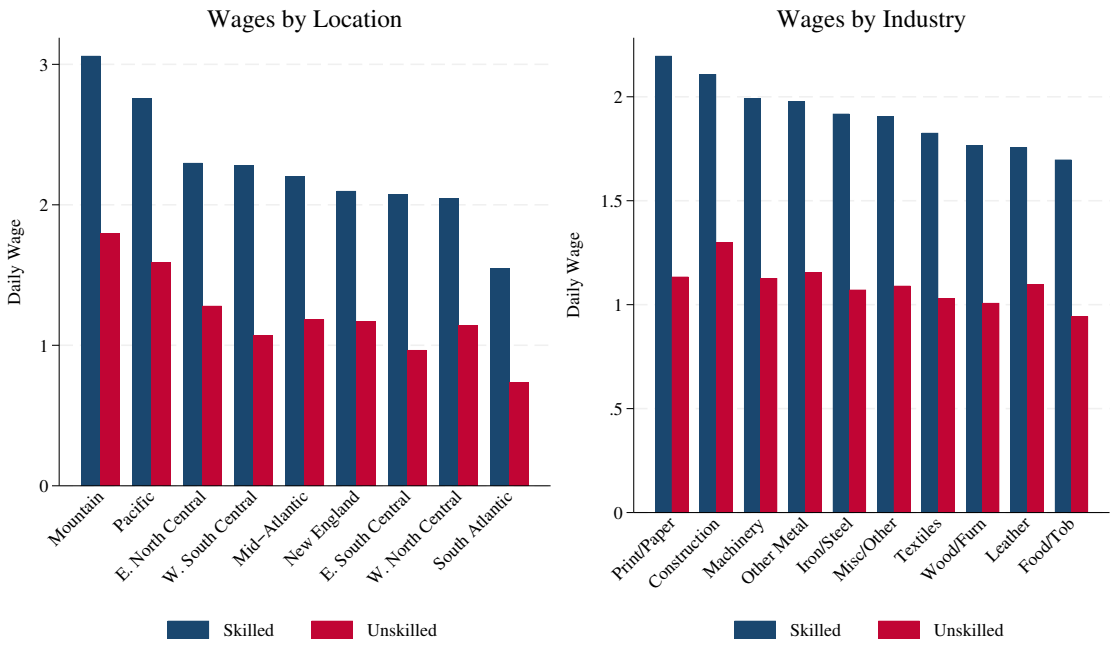


Figure A2: Census of Manufactures firms, 1850–1880

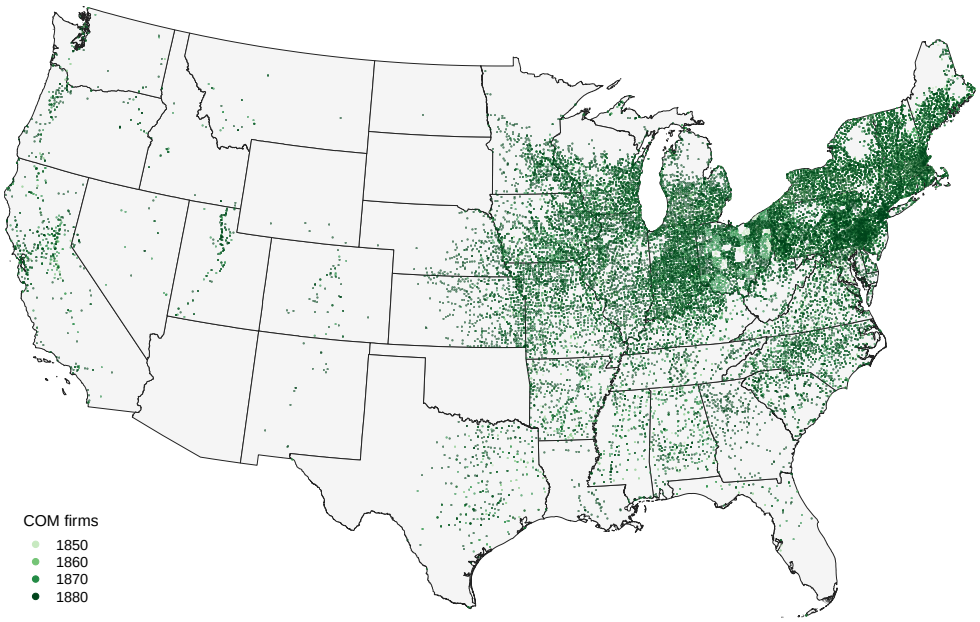


Figure A3: Skilled manufacturing wages, 1880

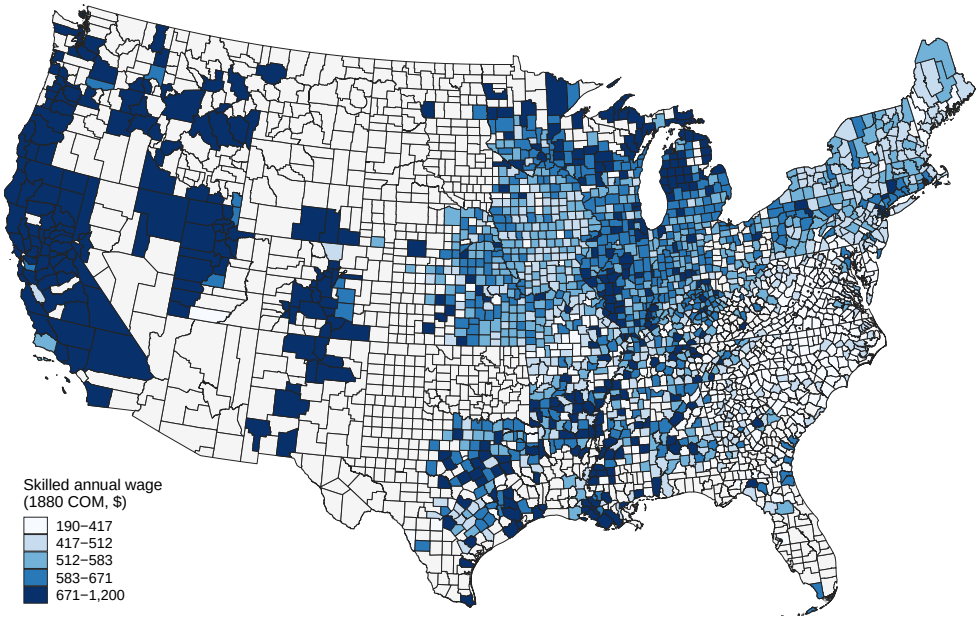


Figure A4: Unskilled manufacturing wages, 1880

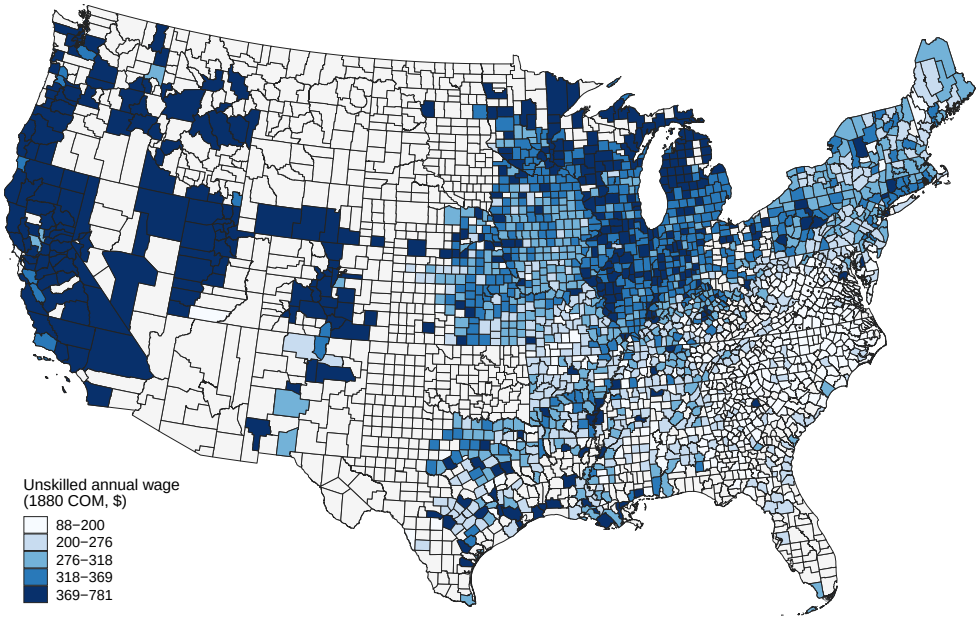


Figure A5: Census Place Project places, 1850–1880

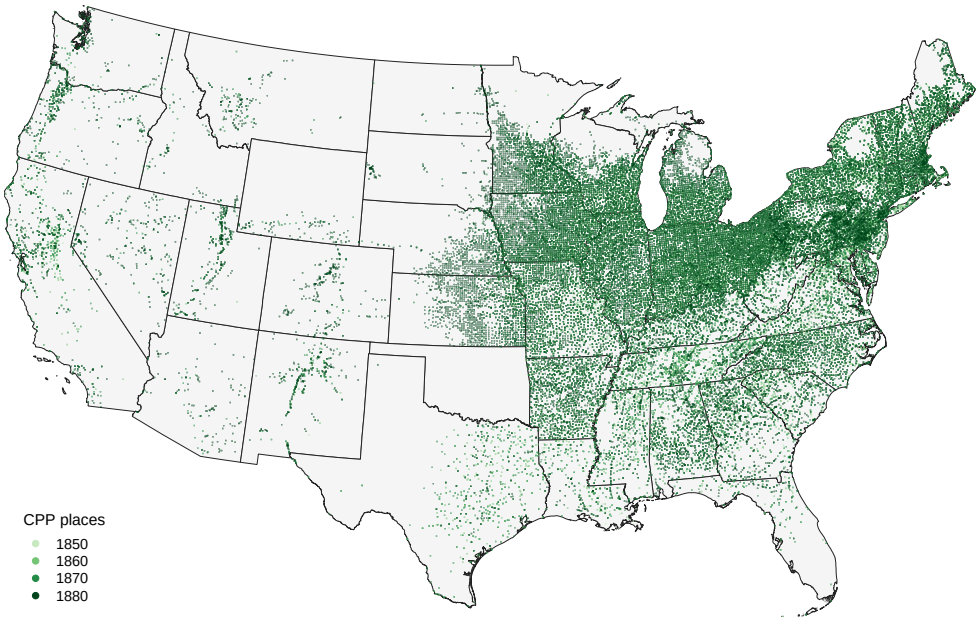


Table A3: Summary statistics for Cost of Living Survey households, 1888–1890

	N	Mean	Median	Std. dev.	Min	Max
Native-born HH head	6,627	0.54	1	0.50	0	1
Wife is working	6,536	0.07	0	0.25	0	1
# of Children	6,627	2.81	3	2.13	0	13
Family size	6,627	4.80	5	2.12	1	15
Husband age	6,627	39.43	38	11.01	18	99
Wife age	6,627	35.54	34	11.91	0	99
Husband's income	6,627	589.34	512	331.81	0	4,986
Wife's income	6,627	12.69	0	57.24	0	886
Children's income	6,627	107.49	0	230.84	0	1,989
Boarder income	6,627	35.98	0	107.82	0	1,828
Other income	6,627	15.57	0	62.15	0	1,330
Total income	6,627	761.07	665	372.44	93	4,986
Avg local income	6,627	591.61	625	130.98	305	1,074

Figure A6: COLS income distribution

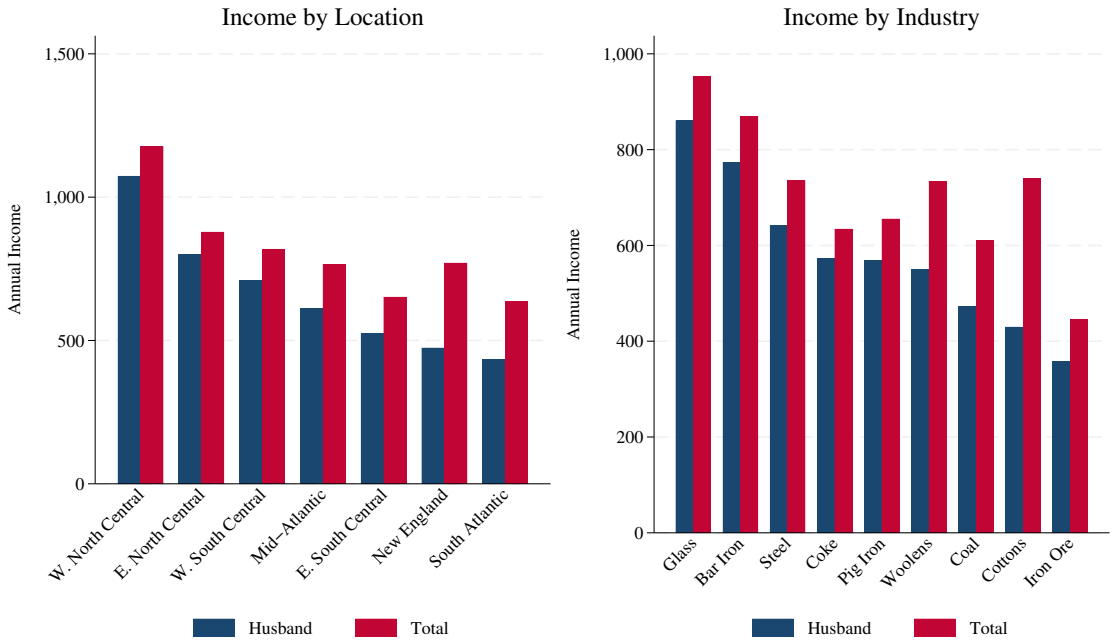


Figure A7: Department of Labor cities



Table A4: Summary statistics for Historical Labor Statistics Project workers, 1879–1903

	N	Mean	Median	Std. dev.	Min	Max
Survey year	64,855	1894.42	1,894	4.55	1,879	1,903
Age in years	59,191	28.42	26	10.85	8	85
Skilled	64,782	0.20	0	0.40	0	1
Adult male	61,812	0.67	1	0.47	0	1
Adult female	61,812	0.29	0	0.45	0	1
Child	61,812	0.04	0	0.20	0	1
County FIPS coded	64,855	0.88	1	0.33	0	1
Daily wage, 1880 dollars	62,928	1.62	1	1.02	0	15
Weekly wage, 1880 dollars (6-day week)	62,930	9.74	8	6.14	0	88
Own annual earnings, 1880 dollars	64,250	438.44	372	302.42	1	4,556
Family annual earnings, 1880 dollars	7,063	456.74	432	255.71	2	3,385

Table A6: Baseline concordance of COLS households and COM titles

COM title	COLS industry	Assignment
Auxiliary trades, any COLS industry		
carpentering	Any	Occupation contains "Carpenter"
blacksmithing	Any	Occupation contains "Blacksmith"
painting	Any	Occupation contains "Painter"
masonry, brick and stone	Any	Occupation contains "Mason" or "Bricklayer"
millwrighting	Any	Occupation contains "Millwright"
machinery, other	Any	Occupation contains "Machinist"
steam engines and boilers	Any	Occupation contains "Engineer" or "Boiler"
patterns and models	Any	Occupation contains "Pattern maker"
cooperage	Any	Occupation contains "Cooper"
boxes, packing	Any	Occupation contains "Boxmaker"
Product assignment for households not in an auxiliary trade		
cotton thread and yarn	Cottons	Occupation contains "Yarn" or "Thread"
cotton goods	Cottons	All other households
wool-carding and cloth-dressing	Woolens	Occupation contains "Card" or "Dresser"
worsted goods	Woolens	Occupation contains "Worsted," and not "Card" or "Dresser"
woolen goods	Woolens	All other households
looking-glass and picture frames	Glass	Occupation contains "Looking-glass"
bottles	Glass	Occupation contains "Bottle," and not "Looking-glass"
glass, n.s.	Glass	All other households
iron and steel, pig	Pig Iron	All households
iron and steel, forged and wrought	Bar Iron	All households
iron and steel, general	Steel	All households
mining, coal	Coal	All households
mining, iron	Iron Ore	All households
coke	Coke	All households

Auxiliary trades are assigned first, so a carpenter in a cotton mill is carpentering. The yarn, thread, worsted, carding, dressing, bottle, and looking-glass rules are applied after those trades and take precedence when an

occupation string matches both. Households that remain "other" are dropped.

Table A7: Baseline concordance of HLSP records and COM titles

COM title	Source	Assignment
Survey industry		
furniture, other	Survey: furniture	All records
iron and steel, forged and wrought	Survey: agricultural implements and iron	All records
cotton goods	Survey: textiles	Industry string contains "cotton," or is blank
woolen goods	Survey: textiles	Industry string contains "wool"
hosiery and knit goods	Survey: textiles	Industry string contains "knit" or "hosier"
masonry, brick and stone	Survey: construction	Occupation contains "brick," "mason," or "stone"
carpentering	Survey: construction	Occupation contains "carpenter"
painting	Survey: construction	Occupation contains "paint"
carpentering	Survey: construction	All other construction records
mining, coal	Survey: stone and clay	Occupation contains "coal"
masonry, brick and stone	Survey: stone and clay	All other stone-and-clay records
Industry string, if still unassigned		
carpentering	Industry string	Contains "carpenter"
blacksmithing	Industry string	Contains "blacksmith"
painting	Industry string	Contains "paint"
paints	Industry string	Recode painting when the string contains "paints" or "varnish" and not "painting"
masonry, brick and stone	Industry string	Contains "mason," "brick," "stone," "potter," "plaster," "stucco," "cement," "marble," or "lime"
machinery, other	Industry string	Contains "foundry" or "machine"
steam engines and boilers	Industry string	Contains "boiler"
iron and steel, forged and wrought	Industry string	Contains "tool," "bolt," "hardware," or "cutlery"
iron and steel, general	Industry string	Contains "iron," "core mak," "plumb," "gasfitt," or "gas fitt"
iron and steel, pig	Industry string	Contains "pig iron"
cooperage	Industry string	Contains "cooper"
boxes, packing	Industry string	Contains "box"
cotton goods	Industry string	Contains "cotton"
woolen goods	Industry string	Contains "wool" or "worsted"
hosiery and knit goods	Industry string	Contains "knit" or "hosier"
boots and shoes	Industry string	Contains "boot" or "shoe"
glass, n.s.	Industry string	Contains "glass"
bottles	Industry string and occupation	Industry string contains "glass" and occupation contains "bottle"
furniture, other	Industry string	Contains "furniture"
printing, job	Industry string	Contains "print," "lithograph," or "bookbind"
bread and bakery products	Industry string	Contains "baker" or "bread"
tobacco, general	Industry string	Contains "cigar" or "tobacco"
lumber, sawed	Industry string	Contains "lumber"
saddlery and harness	Industry string	Contains "saddl" or "harness"
carriages and wagons	Industry string	Contains "carriage," "wagon," or "bugg"
clothing, men's	Industry string	Contains "clothing," "shirt," or "tailor"

Continued on next page

Table A7, continued

COM title	Source	Assignment
flour and grist mills	Industry string	Contains "flour" or "grist"
drugs, chemicals, and medicines	Industry string	Contains "drug," "chemical," or "medicine"
slaughter and meat packing	Industry string	Contains "slaughter," "meat," "butcher," or "packing house"
mining, coal	Industry string	Contains "coal" and not "charcoal"
mining, iron	Industry string	Contains "iron ore" or "iron min"
Occupation; these override the survey and the industry string		
carpentering	Occupation	Contains "carpenter"
blacksmithing	Occupation	Contains "blacksmith" or "horseshoer"
painting	Occupation	Contains "painter"
masonry, brick and stone	Occupation	Contains "mason," "bricklayer," "stone cutter," "stonecutter," or "marble"
millwrighting	Occupation	Contains "millwright"
machinery, other	Occupation	Contains "machinist"
steam engines and boilers	Occupation	Contains "engineer" or "boiler"
patterns and models	Occupation	Contains "pattern"
cooperage	Occupation	Contains "cooper"
boxes, packing	Occupation	Contains "box" and "maker"
Occupation, if still unassigned		
carpentering	Occupation	Contains "cabinet"
saddlery and harness	Occupation	Contains "harness" or "saddl"
masonry, brick and stone	Occupation	Contains "plaster" or "quarry"
iron and steel, general	Occupation	Contains "molder," "moulder," "foundry," or "plumber"
steam engines and boilers	Occupation	Contains "tinner" or "tinsmith"
printing, job	Occupation	Contains "printer," "compositor," "pressman," "press man," or "lithograph"
bread and bakery products	Occupation	Contains "baker"
slaughter and meat packing	Occupation	Contains "butcher" or "meat"
flour and grist mills	Occupation	Contains "miller"
boots and shoes	Occupation	Contains "shoe"
mining, coal	Occupation	Contains "coal miner" or "coal-miner," or contains "miner" and not "iron"
mining, iron	Occupation	Contains "miner" and "iron"
drugs, chemicals, and medicines	Occupation	Contains "druggist" or "pharmacist"
furniture, other	Occupation	Contains "upholster"

Within the survey block, a later row overrides an earlier one when both match. The industry-string block fills records that are still unassigned, except the bottles rule, which applies whenever the industry string contains "glass." The first occupation block then overrides every earlier assignment. Records that remain "other" are dropped.

Table A8: Occupations coded skilled and unskilled

Group	Occupations
Skilled (mechanic / craft)	Accountants And Auditors, Actors And Actresses, Agricultural Sciences-Professors And Instructors, Agricultural Scientists, Architects, Artists And Art Teachers, Athletes, Attendants, Professional And Personal Service (Nec), Authors, Bakers, Bank Tellers, Biological Sciences-Professors And Instructors, Biological Scientists, Blacksmiths, Boilermakers, Bookkeepers, Bricklayers And Masons Apprentice, Brickmasons,Stonemasons, And Tile Setters, Cabinetmakers, Carpenters, Cashiers, Chemical-Engineers, Chemistry-Professors And Instructors, Chemists, Chiropractors, Civil-Engineers, Clergymen, College Presidents And Deans, Compositors And Typesetters, Conductors, Bus And Street Railway, Conductors, Railroad, Dentists, Designers, Draftsmen, Editors And Reporters, Electrical-Engineers, Electricians, Engineering-Professors And Instructors, Engineers (Nec), Engravers, Except Photoengravers, Farm Managers, Floormen And Floor Managers, Store, Foremen (Nec), Forgemen And Hammermen, Funeral Directors And Embalmers, Furriers, Geology And Geophysics-Professors And Instructors, Glaziers, Heaters, Metal, Lawyers And Judges, Librarians, Locomotive Engineers, Loom Fixers, Machinists, Managers And Superintendants, Building, Managers, Officials, And Proprietors (Nec), Mathematics-Professors And Instructors, Mechanical-Engineers, Medical Sciences-Professors And Instructors, Metallurgical, Metallurgists-Engineers, Millwrights, Mining-Engineers, Misc Social Scientists, Misc. Natural Scientists, Molders, Metal, Musicians And Music Teachers, Natural Science (Nec)-Professors And Instructors, Non-Scientific Subjects-Professors And Instructors, Nurses, Professional, Nurses, Student Professional, Officers, Pilots, Pursers And Engineers, Ship, Officials And Administratators (Nec), Public Administration, Officials, Lodge, Society, Union, Etc, Painters, Construction And Maintenance, Painters, Except Construction Or Maintenance, Pattern And Model Makers, Except Paper, Pharmacists, Photographers, Physicians And Surgeons, Physics-Professors And Instructors, Plasterers, Plumbers And Pipe Fitters, Postmasters, Professional, Technical And Kindred Workers (Nec), Psychology-Professors And Instructors, Rollers And Roll Hands, Metal, Social Sciences (Nec)-Professors And Instructors, Sports Instructors And Officials, Stationary Engineers, Statisticians And Actuaries, Stone Cutters And Stone Carvers, Structural Metal Workers, Subject Not Specified-Professors And Instructors, Surveyors, Tailors And Tailoresses, Teachers (N.E.C.), Tool Makers, And Die Makers And Setters, Upholsterers, Veterinarians, Weavers, Textile

Continued on next page

Table A8, continued

Group	Occupations
Unskilled (ordinary laborer and residual)	Advertising Agents And Salesmen, Agents (Nec), Airplane Pilots And Navigators, Apprentices, Building Trades (Nec), Apprentices, Metalworking Trades (Nec), Apprentices, Other Specified Trades, Apprentices, Printing Trades, Apprentices, Trade Not Specified, Asbestos And Insulation Workers, Attendants And Assistants, Library, Attendants, Hospital And Other Institution, Attendants, Physicians And Dentists Office, Attendants, Recreation And Amusement, Auctioneers, Baggagemen, Transportation, Barbers, Beauticians, And Manicurists, Bartenders, Blasters And Powdermen, Boarding And Lodging House Keepers, Boatmen, Canalmen, And Lock Keepers, Bookbinders, Bootblacks, Brakemen, Railroad, Bus Drivers, Buyers And Dept Heads, Store, Buyers And Shippers, Farm Products, Carpenters Apprentice, Cement And Concrete Finishers, Chainmen, Rodmen, And Axmen, Surveying, Charwomen And Cleaners, Clerical And Kindred Workers (N.E.C.), Collectors, Bill And Account, Cooks, Except Private Household, Counter And Fountain Workers, Craftsmen And Kindred Workers (Nec), Cranemen,Derrickmen, And Hoistmen, Credit Men, Dancers And Dancing Teachers, Decorators And Window Dressers, Deliverymen And Routemen, Dispatchers And Starters, Vehicle, Dressmakers And Seamstresses, Except Factory, Dyers, Electricians Apprentice, Electrotypers And Stereotypers, Elevator Operators, Entertainers (Nec), Excavating, Grading, And Road Machinery Operators, Express Messengers And Railway Mail Clerks, Farm And Home Management Advisors, Farm Foremen, Farm Laborers, Wage Workers, Farm Service Laborers, Self-Employed, Farmers (Owners And Tenants), Filers, Grinders, And Polishers, Metal, Firemen, Fire Protection, Fishermen And Oystermen, Foresters And Conservationists, Fruit, Nut, And Vegetable Graders, And Packers, Except Facto, Furnacemen, Smeltermen And Pourers, Gardeners, Except Farm And Groundskeepers, Geologists And Geophysicists, Guards, Watchmen, And Doorkeepers, Heat Treaters, Annealers, Temperers, Housekeepers And Stewards, Except Private Household, Housekeepers, Private Household, Hucksters And Peddlers, Inspectors (Nec), Inspectors, Public Administration, Inspectors, Scalers, And Graders Log And Lumber, Insurance Agents And Brokers, Janitors And Sextons, Jewelers, Watchmakers, Goldsmiths, And Silversmiths, Job Setters, Metal, Laborers (Nec), Laundresses, Private Household, Laundry And Dry Cleaning Operatives, Linemen And Servicemen, Telegraph, Telephone, And Power, Locomotive Firemen, Longshoremen And Stevedores, Lumbermen, Raftsmen, And Woodchoppers, Machinists And Toolmakers Apprentice, Mail Carriers, Marshals And Constables, Mathematicians, Meat Cutters, Except Slaughter And Packing House, Mechanics And Repairmen (Nec), Mechanics, Except Auto Apprentice, Medical And Dental-Technicians, Members Of The Armed Services, Messengers And Office Boys, Midwives, Millers, Grain, Flour, Feed, Etc, Milliners, Mine Operatives And Laborers, Motion Picture Projectionists, Motormen, Mine, Factory, Logging Camp, Etc, Motormen, Street, Subway, And Elevated Railway, Newsboys, Office Machine Operators, Oilers And Greaser, Except Auto, Operative And Kindred Workers (Nec), Opticians And Lens Grinders And Polishers, Paperhangers, Photoengravers And Lithographers, Photographic Process Workers, Physicists, Piano And Organ Tuners And Repairmen, Plumbers And Pipe Fitters Apprentice, Policemen And Detectives, Porters, Power Station Operators, Practical Nurses, Pressmen And Plate Printers, Printing, Private Household Workers (Nec), Purchasing Agents And Buyers (Nec), Radio Operators, Railroad And Car Shop-Mechanics And Repairmen, Real Estate Agents And Brokers, Recreation And Group Workers, Religious Workers, Roofers And Slaters, Sailors And Deck Hands, Salesmen And Sales Clerks (Nec), Sawyers, Service Workers, Except Private Household (Nec), Sheriffs And Bailiffs, Shipping And Receiving Clerks, Shoemakers And Repairers, Except Factory, Social And Welfare Workers, Except Group, Spinners, Textile, Stationary Firemen, Stenographers, Typists, And Secretaries, Stock And Bond Salesmen, Switchmen, Railroad, Taxicab Drivers And Chauffeurs, Teamsters, Technicians (Nec), Telegraph Messengers, Telegraph Operators, Telephone Operators, Testing-Technicians, Therapists And Healers (Nec), Ticket, Station, And Express Agents, Tinsmiths, Coppersmiths, And Sheet Metal Workers, Truck And Tractor Drivers, Ushers, Recreation And Amusement, Waiters And Waitresses, Watchmen (Crossing) And Bridge Tenders, Welders And Flame Cutters

Skill follows the unified COLS-to-COP-to-HLSP mapping: COLS craft titles (painters, weavers, masons, glass blowers/cutters, puddlers, heaters, etc.) are skilled in every source; remaining named professional/craft titles are skilled; named unskilled titles and residual occupations are unskilled. Non-occupation codes 979–999 are omitted.

Table A9: Wage coverage of the 1880 labor force

Source	Workers	Share
Manufacturing, spatial	1,271,415	7.63%
Manufacturing, county fallback	641,938	3.85%
Manufacturing, state fallback	536,732	3.22%
Manufacturing, other fallback	124,199	0.74%
GPIH, state	10,580,489	63.46%
GPIH, region	231,522	1.39%
GPIH, national	170,183	1.02%
Uncovered	3,115,254	18.69%
Total	16,671,732	100.00%

A manufacturing wage is assigned first. State, region, and national quotes cover workers with no manufacturing wage. Of the state row, 9,986,633 workers have a reported quote and 593,856 are rank-filled. There is no county quote.

Table A10: Wage sources in the wage component of the composite index

Source	Year	Geographic scope	1880 workers (share)
COM	1880	Spatial, county	2,785,532 (16.71%)
Anchors and chains, Awnings and tents, Bagging, flax, hemp, and jute, Bags, paper, Baking and yeast powders, Baskets, Bells, Belting and hose, Blacking and other polishes, Blacksmithing, Bleaching, dyeing, and cleaning, Blocks, pumps, masts, and spars, Boat building, Bolts, nuts, washers, and rivets, Bone-, ivory-, and lamp-black, Boots and shoes, Boxes, cigar and tobacco, Boxes, fancy and paper, Boxes, other, Boxes, packing, Brassware and bells, Bread and bakery products, Brick, Brick, stone, and tile, n.s., Bronze, Butter and cheese, Cabinet ware, Canning and preserving, Carpentering, Carpets, n.s., Carpets, rag, Carriages and wagons, n.s., Cement, Chemical pigments, other, Chemicals, other, Chocolate, Clocks, Clocks and watches, n.s., Clothing, children's, Clothing, general, Clothing, men's, Clothing, n.s., Clothing, women's, Coffee and spices, roasted and ground, Confectionery, Construction, n.s., Construction, other, Copper, Copper ware, Cordage and twine, Cotton goods, Crucibles, Cutlery, edge tools, and axes, Drain pipe, Drugs, chemicals, and medicines, Dyestuffs and extracts, Earthenware, Electrical, telegraph, and telephone apparatus, Electroplating, Envelopes, Explosives and fireworks, Fertilizers, Files, Fish and oysters, canned, Flags and banners, Flax, dressed, Flour and grist mills, Food products, other, Fruits and vegetables, canned and preserved, Fuel, kerosene and camphene, Fuel, other, Furniture, other, Furs, Gas and lamp fixtures, Glass, n.s., Gloves and mittens, Glue, Gold and silver leaf and foil, Gold and silver refining, Gun- and lock-smithing, n.s., Gunpowder, Gunsmithing, Handles, wooden, Hardware, n.s., Hardware, saddlery, Hats and caps, Heating apparatus, Hones and whetstones, Hosiery and knit goods, Ink, Iron and steel products, other, Iron and steel, forged and wrought, Iron and steel, general, Iron and steel, other, Iron and steel, pig, Japanned ware, Jewelry, Kaolin and ground earths, Kindling wood, Lead, Leather products, other, Leather, curried, Leather, dressed skins, Leather, morocco, Leather, n.s., Leather, other, Leather, tanned, Lime, Linen and linen goods, Liquors and beverages, other, Liquors, distilled, Liquors, malt, Liquors, rectified, Liquors, vinous, Locksmithing and bellhanging, Lumber, sawed, Machinery, cotton and woolen, Machinery, n.s., Machinery, other, Malt, Masonry, brick and stone, Matches, Mats and rugs, Meters, Millinery, Millstones, Millwrighting, Mineral and soda waters, Mining, coal, Mining, iron, Mining, n.s., Mining, other, Mixed textiles, Nails and spikes, Non-metal minerals, other, Oilcloth, Oils, lubricating, Other metal products, Painting, Paints, Paper, Paper collars and cuffs, Paper, other, Paperhangings, Perfumery and cosmetics, Photography, Pickles, preserves, and sauces, Plaster, Plastering, Pocket-books, Printing, job, Pumps, Quarrying, Rubber and elastic goods, Saddlery and harness, Sails, Sand and emery paper and cloth, Saws, Scales and balances, Screws, Sewing machines and attachments, Ship and boat building, n.s., Ship and boat building, other, Ship building, Shoddy, Silk and silk goods, Silverware, Slaughter and meat packing, Smelting and refining, other, Soap and candles, Springs, Starch, Stationery and school supplies, Steam engines and boilers, Stone ware, Stoves and ranges, Straw goods, Sugar and molasses, Tailoring, Textile products, other, Tin ware, Tools, Trunks, carpet bags, and valises, Umbrellas, whips, and canes, Unclassified, Vinegar, Washing machines and dryers, Watch and clock materials, Watch and clock repairing, Watches, White lead, Willow ware, Wire, Wirework, Wood products, other, Wood, turned and carved, Wooden ware, Wool-carding and cloth-dressing, Zinc			
GPIH	1870	State, region, US	10,004,584 (60.02%)
Farm laborers, Laborers (n.e.c.), Private household workers			
GPIH	1870	State, region	64,033 (0.38%)

Continued on next page

Table A10, continued

Source	Year	Geographic scope	1880 workers (share)
Clergy			
GPIH	1870	State, US	256,624 (1.54%)
Mine operatives			
GPIH	1870	Region, US	862,445 (5.17%)
Blacksmiths, Carpenters, Machinists, Painters, Sailors, deck hands, and boatmen			
GPIH	1870	Region	295,601 (1.77%)
Boilermakers, Brickmasons and stonemasons, Cabinetmakers, Compositors, Molders, Pattern and model makers, Plumbers, Stone cutters, Teamsters			
GPIH	1870	US	57,015 (0.34%)
Stationary engineers			
GPIH	1860	State, region, US	157,459 (0.94%)
Clerks			
GPIH	1860	State, region	77,736 (0.47%)
Lawyers and judges, Lumbermen			
GPIH	1860	US	6,742 (0.04%)
Civil engineers			
GPIH	1850, 1860	State	22,486 (0.13%)
Teachers, Professors, Physicians, Librarians			

COM is the 1880 Census of Manufactures. A person uses the employment-weighted mean of firm payroll per worker in that industry within 15 miles of the Census Place Project place when any such firm is present, and otherwise the county mean. Annual pay is expressed in United States 1880 dollars with the Coelho–Shepherd regional price index and converted to a daily rate by dividing by 300. When a Bayes-shrunk place × occupation × demographic wage exists, that daily rate is used. GPIH is the Lindert–Williamson quote. Cash rates are converted to a daily rate (weekly/6, monthly/26, annual/300; hourly times reported hours, or 10 hours when hours are missing) and deflated with the same index. Quotes for 1850 use the 1851 Coelho–Shepherd level. In-kind board is excluded. Rows group occupations by the geography on the sheet. A person receives a 1870 quote when one exists, then 1860, then 1850. Within a quote year the match is the state, then the Coelho region, then the United States. Own-account farmers and unpaid family farm workers receive the farm labor wage, and that occupation's count includes farmers. Worker counts are people in the 1880 full count with a reported occupation (OCC1950 codes 1–979; N = 16,669,500). Shares are of that workforce. The manufacturing count sums the industry titles, and a person can be counted in both panels. Clerks, lawyers and judges, teachers, professors, civil engineers, lumbermen, physicians, and librarians have no 1870 quote; they receive an 1860 or 1850 quote when one exists for their geography.

Table A11: Spatial wage robustness: rank-rank correlations with external wages

	DOL	HLSP manufacturing	HLSP all workers
Baseline	0.606 (162)	0.756 (375)	0.851 (910)
Alternative Distance Thresholds			
10-mile radius	0.610 (162)	0.652 (350)	0.817 (907)
5-mile radius	0.588 (162)	0.708 (311)	0.846 (898)
Inverse-distance weights	0.603 (162)	0.744 (375)	0.849 (910)

Continued on next page

Table A11, continued

	DOL	HLSP manufacturing	HLSP all workers
Alternative Firm Weighting			
Unweighted firms	0.601 (162)	0.756 (375)	0.845 (910)
Maximum employment	0.594 (162)	0.757 (375)	0.845 (910)
Wage bill	0.602 (162)	0.755 (374)	0.846 (909)
Alternative Individual Weighting			
Unweighted cells	0.627 (162)	0.763 (375)	0.858 (910)
Manufacturing workers	0.638 (162)	0.758 (375)	0.855 (910)
Census population	0.638 (162)	0.761 (375)	0.853 (910)
Ridge constant 0.5	0.606 (162)	0.756 (375)	0.851 (910)
Occupation scores weighted by population	0.606 (162)	0.756 (375)	0.851 (910)
County Wages (not CPP-Spatial)			
County, demographic group, skill, and industry	0.674 (151)	0.791 (235)	0.810 (606)
Without Various Dimensions			
No skill	0.506 (149)	0.463 (337)	0.733 (913)
No demographic group	0.638 (163)	0.730 (359)	0.866 (824)
No industry	0.833 (42)	0.731 (477)	0.732 (519)
Decimal corrections			
Divide by 100 when wage exceeds 10 times the county median	0.560 (163)	0.363 (374)	0.758 (922)
Divide by 100 when wage exceeds 5 times the county median	0.550 (163)	0.382 (374)	0.766 (922)
Divide by 100 when wage exceeds 10 times the industry median	0.581 (163)	0.596 (374)	0.822 (922)
Divide by 100 when wage exceeds 5 times the industry median	0.572 (163)	0.596 (374)	0.822 (922)
Winsorization			
No winsor	0.571 (163)	0.592 (374)	0.820 (922)
Winsor at the 5th and 95th percentiles	0.578 (163)	0.599 (374)	0.830 (922)
Winsor within industry	0.572 (163)	0.596 (374)	0.822 (922)
Winsor within state	0.573 (163)	0.596 (374)	0.822 (922)
Bayes shrinkage			
Shrunk toward occupation, region, and demographic group	0.642 (162)	0.743 (511)	0.849 (882)

Each entry is the rank-rank correlation of that spatial-wage variant with the external wage, on the same cells as the main validation figure. The number of cells is in parentheses. DOL pools the 1870 and 1880 city × skill × industry wages and is unweighted. Both HLSP columns are worker-weighted county × demographic group × skill × industry cells. Manufacturing uses the spatial wage. All workers adds agriculture, domestic service, hack and bus, teachers, clerks, and laborers at the GPIH quote, which does not change across rows. A variant replaces the 1880 manufacturing wage only. A row that drops demographic group, skill, or industry is matched after collapsing that dimension. The Bayes row assigns each 1880 worker with a panel wage the daily wage shrunk toward the occupation × census region × demographic group mean (white men, non-white men, and women). A child keeps the shrunk wage of their place, occupation, and race-sex group.

Table A12: Inequality implied by standardized SES scores, 1880

	Manufacturing workers			All workers		
	Individual	Region	State	Individual	Region	State
Composite SES	0.47	0.11	0.15	0.44	0.20	0.23
Equal weights	0.58	0.07	0.10	0.46	0.19	0.22
Spatial wage	0.54	0.14	0.19	0.67	0.24	0.28
Literacy share	0.20	0.05	0.07	0.28	0.16	0.18
Average wealth	0.67	0.05	0.10	0.60	0.16	0.23
Duncan SEI	0.71	0.02	0.09	0.71	0.15	0.19
Collins-Zimran wealth	0.67	0.05	0.13	0.63	0.20	0.28
Occupational education	0.61	0.02	0.07	0.64	0.08	0.13
Occscore	0.59	0.02	0.07	0.60	0.17	0.19
LIDO	0.58	0.05	0.13	0.59	0.20	0.26
Siegel prestige	0.57	0.04	0.07	0.55	0.06	0.08
Nam-Powers-Boyd	0.54	0.02	0.06	0.64	0.18	0.21
Occupational earnings	0.51	0.03	0.08	0.59	0.18	0.21
Adjusted Song	0.45	0.15	0.18	0.54	0.23	0.25
Song	0.39	0.06	0.09	0.55	0.11	0.14

Each entry is a Gini coefficient. Scores are standardized to mean 0 and standard deviation 1 and then exponentiated for common scale. Region and state columns assign each worker the mean transformed score of their census region or state and then compute the Gini. Manufacturing workers have a Census of Manufactures wage. All workers have that wage or a GPIH quote. Composite SES is observed for 2,565,824 manufacturing workers and 13,554,037 workers overall. LIDO is observed for 2,212,916 manufacturing workers and 9,880,595 workers overall.

Table A13: All-worker Gini coefficients, 1850–1880

	1850	1860	1870	1880	1880−1850
Composite SES	0.31	0.38	0.51	0.51	0.20
Equal weights	0.30	0.57	0.52	0.50	0.20
Spatial wage	0.41	0.50	0.52	0.57	0.16
Average wealth	0.40	0.75	0.61	0.60	0.19
Occupational education	0.46	0.50	0.46	0.50	0.04
LIDO	0.55	0.61	0.61	0.60	0.06
Occupational earnings	0.54	0.59	0.59	0.60	0.06
Nam-Powers-Boyd	0.56	0.63	0.64	0.65	0.09
Occscore	0.49	0.58	0.58	0.60	0.11
Adjusted Song	0.51	0.57	0.62	0.63	0.11
Duncan SEI	0.55	0.62	0.63	0.67	0.12
Collins-Zimran wealth	0.44	0.51	0.54	0.57	0.13
Siegel prestige	0.40	0.51	0.54	0.54	0.14
Song	0.40	0.49	0.55	0.55	0.15

Figure A8: Father-son persistence: Alternative specifications

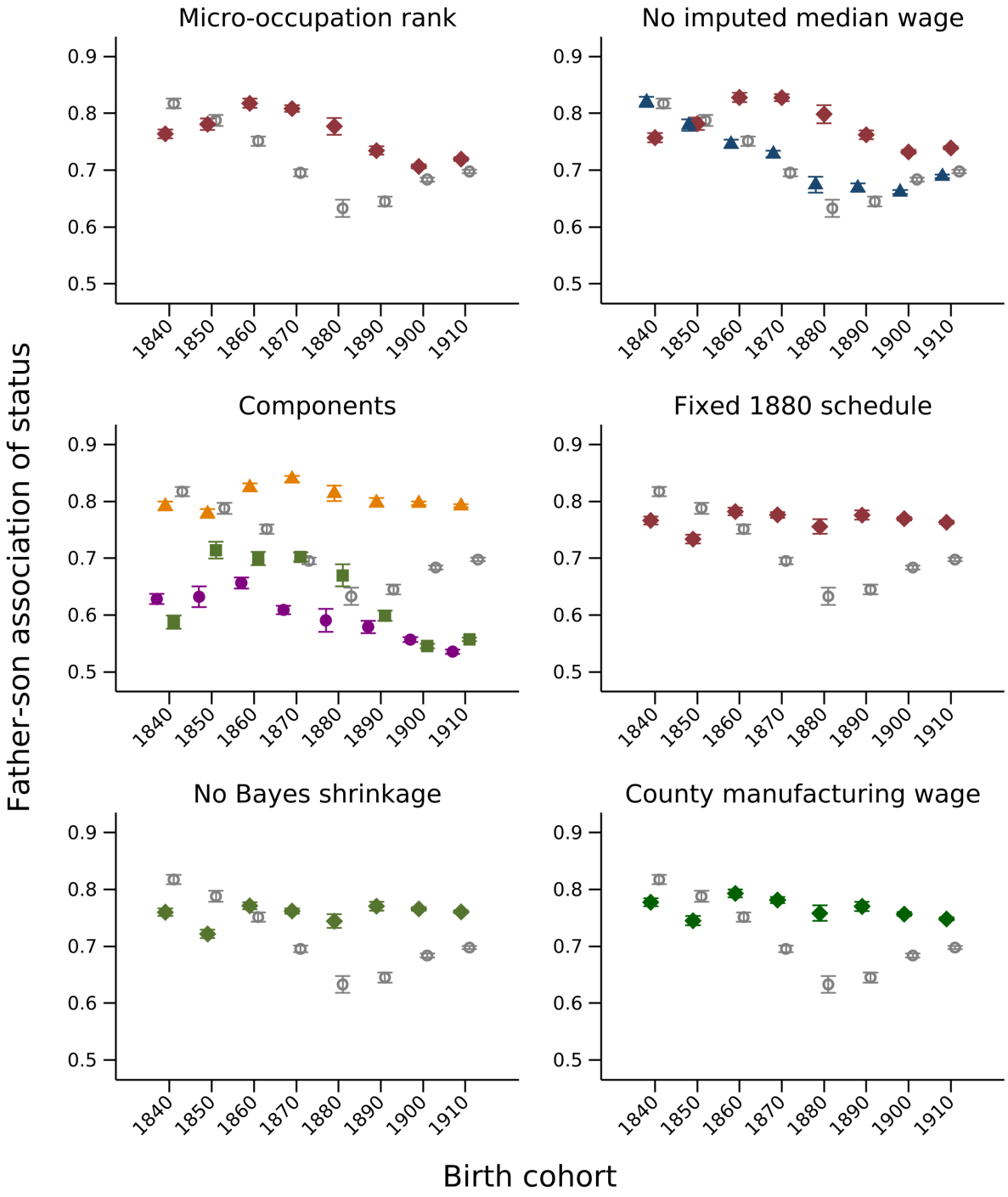


Figure A9: Father-son persistence: Alternative SES measures

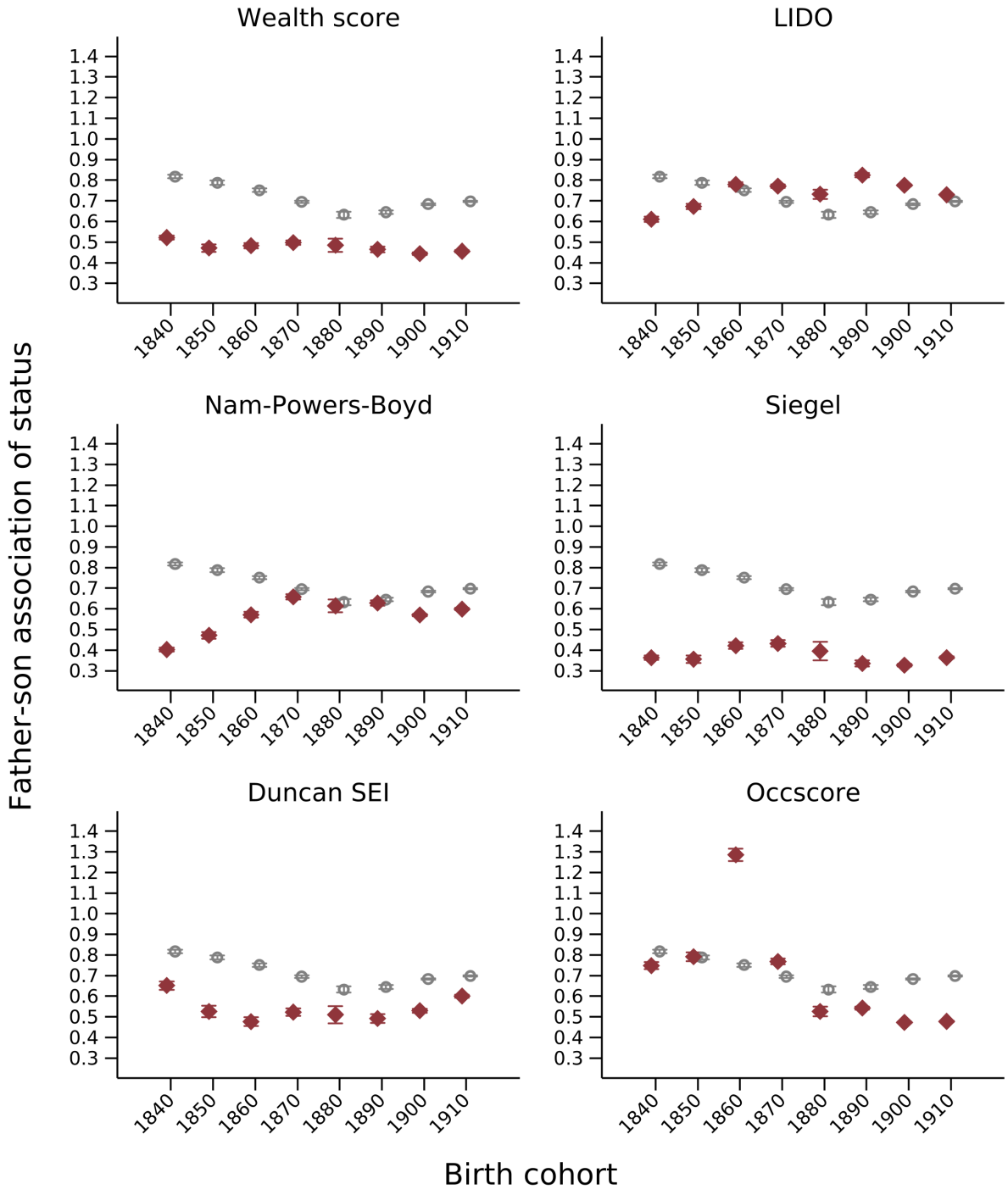


Table A14: Manufacturing 90th-to-10th ratios, 1850–1880

	1850	1860	1870	1880	1880−1850
Composite SES	2.54	4.19	8.73	10.54	8.01
Equal weights	2.06	3.81	6.93	6.28	4.22
Spatial wage	2.89	4.58	7.42	14.21	11.32
Average wealth	1.46	3.47	2.18	2.37	0.91
Nam-Powers-Boyd	18.23	29.04	20.88	16.84	-1.38
Collins-Zimran wealth	3.09	1.79	1.78	1.95	-1.13
Duncan SEI	4.86	5.74	5.74	4.48	-0.39
Occupational earnings	8.41	12.14	8.03	8.20	-0.21
Song	9.99	10.15	10.15	10.15	0.16
Occupational education	1.50	1.71	1.50	2.24	0.74
Adjusted Song	10.11	10.31	10.63	10.90	0.79
Siegel prestige	5.67	6.30	5.67	6.49	0.82
LIDO	4.72	4.72	4.34	6.14	1.41
Occscore	5.86	8.35	4.91	8.35	2.49