# Rising Wage Dispersion Across American Manufacturing Establishments, 1850 - 1880
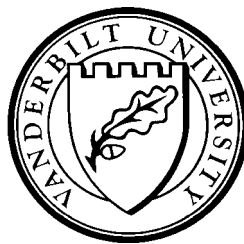
by

Jeremy Atack, Fred Bateman, and Robert A. Margo

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

RISING WAGE DISPERSION ACROSS AMERICAN MANUFACTURING

ESTABLISHMENTS, 1850-1880


by

Jeremy Atack

and

Fred Bateman

and

Robert A. Margo

Revised, September 2000

Jeremy Atack is Professor of Economics, Vanderbilt University, and Research Associate, National Bureau of Economic Research. Fred Bateman is Nicholas A. Beadles Professor of Economics, Terry College of Business, University of Georgia. Robert A. Margo is Professor of Economics, Vanderbilt University; Research Associate, National Bureau of Economic Research; and Visiting Senior Scholar, The Jerome Levy Economics Institute of Bard College. We are grateful to Stanley Engerman, Claudia Goldin, Daniel Hamermesh, Lawrence Katz, and workshop participants at the University of Texas, Harvard University, and the Fourth World Congress of Cliometrics in Montreal, Canada (July 2000) for helpful comments. Corresponding author: Robert A. Margo, The Jerome Levy Economics Institute of Bard College, Blithewood, Annandale-on-Hudson, New York 12504-5000, margo@levy.org

ABSTRACT

1

We use data from the manuscript censuses of manufacturing for 1850, 1860, 1870, and 1880 to study the dispersion of average monthly wages across establishments. We find a marked increased in wage inequality over the period, an increase that cannot be explained by biases in the data or changes in census enumeration procedures. Based on log wage regressions on establishment characteristics (for example, size, capital intensity, use of steam power, and so on) we compute a decomposition of the change in wage inequality between 1850 and 1880. The decomposition reveals that changes in "wage structure" – the regression coefficients and the standard error of the residuals – largely offset each: changes in the coefficients produced a reduction in wage inequality, while residual inequality increased. Most of the rise in wage inequality can be attributed to an increased concentration of employment in large establishments, which paid relatively low wages. We present indirect evidence that the negative effect of size on wages reflected differences in skill composition: workforces in large establishments were less skilled than in small establishments.

Jeremy Atack
Department of Economics
Vanderbilt University
Nashville TN 37235
and NBER
jeremy.atack@vanderbilt.edu

Fred Bateman
Department of Economics
Terry College of Business
University of Georgia
Athens GA 30602
bateman@terry.uga.edu

Robert A. Margo
The Jerome Levy Economics Institute of Bard College
Blithewood
Annandale-on-Hudson NY 12504-5000
and NBER
margo@levy.org

1. Introduction

Changes in the structure of wages are a prominent feature of the economic history of many nations.[1]  As technology advances, the demand for one type of labor may expand relative to another, driving up the relative wage of the former at the expense of the latter.  In a similar manner, differences in factor endowments across regions or countries may alter the relative productivity of, say, unskilled labor relative to skilled labor, or women and children versus adult men (Habakkuk 1962; Goldin and Sokoloff 1982).  Shocks – such as wars or business cycles – may temporarily alter the wage structure, either through changes in relative demands or supplies or government intervention.  Whether shifts in the structure of wages are transitory or persistent depends on the responsiveness of relative supplies of labor of different types to changing wage differentials.

Given its importance, it is surprising how many gaps there are in the documentation of the history of the American wage structure. By and large, research has concentrated on the period from 1940 to the present, in part because of the availability of large, nationally-representative samples containing individual-level information on wages and worker characteristics.  Based on such evidence it is known, for example, that Americans experienced a "Great Compression" between 1940 and 1950 during which wage differentials within and between labor market groups (for example, college versus high school graduates)  narrowed substantially, partly because of changes in the relative demand for unskilled labor associated with World War Two and because of direct government intervention in the wage-setting process (Goldin and Margo 1992).   In the 1970s the returns to college fell and that of labor market experience rose, as a consequence of the entry of the "baby-boom" generation into the labor market.  But, in the 1980s, wage premiums associated with measurable skills increased substantially, as a result of a slowdown in the growth of the relative supply of educated labor, as well as a host of technological and other changes that increased

3

relative demand  (Juhn, Murphy, and Pierce 1993; Katz and Murphy 1992).

Because of the virtual absence of such micro-data prior to 1940, research on the pre-1940 wage structure has relied, either wholly or in part, on evidence on occupational wages, usually summarized in the form of <u>skill differentials</u>: the ratio of wages in one or more "skilled" occupations, relative to wages of "common" or unskilled labor.  For the period 1900-1940 much evidence points to a substantial secular decline in the relative wage of skilled labor (see the references in Williamson and Lindert 1980).  The decline in skill differentials during the first part of the 20th century appears to be explained by rapid growth in the supply of educated labor, a consequence of the "so-called" high school movement.  Even though a variety of technical and other changes in the economy led to a pronounced increase in the relative demand for educated labor after the turn of the century, the supply shift was so great that the returns to education fell, as evidenced by relative declines in the wages of a variety of occupations employing skilled or educated labor, as well as direct evidence for one state (Iowa, see Goldin and Katz 1998, 1999).

For the 19th century there is far less agreement about the course of skill differentials (let alone the causes of any changes that may have occurred).  Until recently, for example, it was widely believed that wages of skilled blue collar workers – artisans -- increased relative to common labor over the period 1800 to 1850 – the so-called "ante-bellum surge" (Williamson and Lindert 1980).  More recent research suggests, however, that no surge in the relative wages of artisans took place before the Civil War; if anything, the skill differential measured in these terms was declining prior to the 1850s (Grosse 1982; Margo and Villaflor 1987; Margo 2000, Table 5B.5).  However, it does appear that the relative wages of white-collar workers – the educated labor of the period – were rising over the 1820 to 1850 period,  possibly because of growth in the relative demand for managerial and related skills associated with early industrialization and a

slowdown in the rate of white-collar immigration to the United States after 1840 (Margo 2000, chs. 5, 7; Cohn 2000).

For the second half of the 19th century the available evidence on trends in skill differentials is conflicting. On the one hand, the ratio of daily pay of engineers to common labor employed in manufacturing, mining, and mechanical establishments increased by approximately 15 percent comparing the 1870s to the 1850s (Coehlo and Shepherd 1976, as reprinted in Williamson and Lindert 1980, p. 306). Similarly, a series of the ratio of artisans-to-laborers pay suggests a substantial rise in the skill differential in the 1860s and 1870s, followed by a decline in the 1880s to an average decade level about 7 percent higher than in the early 1850s (Williamson and Lindert 1980, appendix D, series 3, p. 306). However, another series pertaining to artisans and common labor shows no trend in pay ratios between 1850 and 1890 (Williamson and Lindert 1980, Appendix D, series 5, p. 307). But all of these series are based on data that are known, a priori, to seriously under-represent various industries and locations, or are inherently biased in ways that may compromise their ability to faithfully replicate trends in the broader economy.[2]

The purpose of this paper, however, is not to criticize previous work on 19th century skill differentials, or to present yet more series of skill differentials based on the same (or similar) data used in prior studies. Rather, this paper presents new archival evidence on the wage structure in the United States for a portion of the second half of the 19th century, the period from 1850 to 1880. The analysis is based on recently collected random samples from the manuscript censuses of manufacturing. The unit of observation is the establishment, and the variable examined is the average monthly wages of employees at the establishment, which we call the "establishment wage". Our most important finding is that the degree of inequality in establishment wages increased substantially between 1850 and 1880. Although some of the rise in wage dispersion

may reflect changes in census enumeration procedures over time, we believe that the bulk of the increase is a real historical phenomenon, one that has previously escaped the attention of economic historians.

Having established that a rise in wage inequality occurred, we turn our attention to explaining the increase. We estimate log wage regressions in which the independent variables are characteristics of the workers (for example, percent female) or the establishment (for example, the capital-labor ratio), and use the results of the regression to compute a decomposition analysis of the change in wage dispersion between 1850 and 1880. We find that changes over time in the regression coefficients produced a narrowing, not a widening, of wage inequality. Because a case can be made that (at least some of) the variables in the regression are correlated with the skill composition of establishment workforces, our data provide little support for the view that the skill differential, as it is conventionally defined, was rising between 1850 and 1880. However, such conclusion is misleading about changes in the overall structure of wages across establishments, because residual wage inequality increased, enough to (slightly) offset the compression in the regression coefficients.

Perhaps our most interesting finding, however, is that the majority of the increase in wage inequality can be attributed to an increasing share of employment in very large establishments, which paid lower than average wages. One of the most basic (and pervasive) features of the evolution of manufacturing in the 19th century was the substitution of unskilled "operatives" for artisans in the course of the diffusion of the factory system. As the share of employment in large firms grew, so too did the density of low-wage establishments in the overall wage distribution. While labor historians have frequently alluded to the importance of the "de-skilling" of the manufacturing labor force in these terms (see, for example, Brown and Philips 1986), to our

6

knowledge ours is the first study to measure the quantitative impact on wages at the establishment level. Because manufacturing's share of the aggregate labor force increased after 1850, our findings suggest that, if de-skilling produced a general widening of wage inequality in the aggregate economy, it did so by increasing the relative proportion of less-skilled workers in the aggregate labor force.


2. Evidence on the Distribution of Wages Across Manufacturing Establishments, 1850-1880


Our evidence derives from samples from the decennial manuscript censuses of manufacturing for the years 1850, 1860, 1870 and 1880 collected and computerized by Jeremy Atack and Fred Bateman (1999). The samples are random and nationally representative of the universe of surviving manuscript schedules. Each contains several thousand establishments. The census enumerators were directed to obtain information through personal inquiry of cognizant parties including the value of real and personal estate invested in each enterprise; the labor employed and wages paid; the values of inputs and outputs; the type of power used; and, after 1860, the months of operation. The data pertain to economic activities from June of the year prior to the census through May of the census year.

The earlier censuses differed from the later ones in the wage and employment information they collected. In 1850 data were collected on the total monthly wages paid to male and female labor, and the number of male and female workers employed. Beginning in 1870 the question on monthly wages was dispensed with, and replaced by a question on the total annual wage bill; employees were classified into adult males (over age 16), females (over age15), and children of both sexes; and (as noted earlier) the number of months the firm was in operation was reported. In 1880 questions were added about the average daily wage of skilled and unskilled labor (see section 4) and a more complex

set regarding hours and months of operation. An important limitation of these data – indeed, of all extant 19th century establishment-level data on manufacturing – is that we cannot distinguish between production and non-production workers, unlike in studies based on 20th century data (e.g. Allen 1995; Goldin and Katz 1998).

Our analysis is conducted in terms of average monthly wages at the level of the establishment, the basic unit of observation in the census. In 1850 and 1860 the average monthly wage of the jth establishment ($w_{.j}$) is

$$w_{.j} = \text{(total monthly wages, male + total monthly wages, female)/(total employees)}$$

while in 1870 and 1880, the average monthly wage is

$$w_{.j} = \text{(annual wage bill/months of operation)/(total employees)}$$

Throughout the paper we refer to $w_j$ as the "establishment wage."

Studies of wage inequality using post-1940 data typically impose restrictions on the support of the distribution of wages (Goldin and Margo 1992; Katz and Murphy 1992). Very low wage observations are deleted as outliers while, motivated partly by concerns of privacy, very high-wage observations have been "top-coded".[3] We have followed similar procedures by imposing lower and upper bounds although, in both cases, our motivation is to minimize the impact of severe outliers, which we believe primarily reflect mis-reported or erroneous observations.[4] These upper and lower bounds are indicated in the notes in Table 1 and are maintained throughout the analysis.

Our analysis is conducted in terms of weighted measures of the inequality of the logarithm of

$w_{.j}$ (ln $w_j$) where the weight is the establishment size, as measured by the number of workers employed. Such measures can be interpreted as the "between-group" component of wage inequality at the level of individual workers, where the establishment is the grouping variable. That is, in terms of the distribution of individual earnings, we are replacing each worker's wage with the establishment average, and then computing the measure of inequality.[5]

Panel A of Table 1 shows estimates of three different types of inequality measures: the 90-10 range statistic (the difference between the log wage at the $90^{th}$ versus the $10^{th}$ percentiles); the difference between means of the ninth and first deciles of the log wage distribution, which we refer to as the "9-1 wage gap; and the standard deviation. All of the measures indicate an increase in inequality between 1850 and 1880, that of the 90-10 range statistic being the largest in percentage terms. Although all three measures show an upward drift in the 1850s, much of the overall increase between 1850 and 1880 occurred between 1860 and 1870 – the decade of the Civil War. Later in the paper we consider a possible role for the War in generating a rise in wage inequality. However, it is clear that any immediate impact of the War cannot be the full explanation, because the increase in wage dispersion was sustained into the 1870s – indeed, the rise in the 90-10 range statistic in the 1870s exceeded that of the 1860s.

The data in Panel A are not "controlled" for in any way – that is, we do not adjust for any factors at the establishment level that might have been correlated with the average wage. For the moment, however, we focus on another issue: could the rise in wage inequality be a statistical artifact? As already noted, the samples we analyze are random in design, so that is not the force of the objection. Rather, the issue is whether various inconsistencies or biases inherent in the original data may be responsible for the increase in wage inequality. Although not all of the inconsistencies can be quantified or even identified, we can examine certain sources of inconsistency and bias.

9

One possibility concerns the change over time, noted earlier, in the definition of the monthly wage; recall that, in 1850 and 1860, the monthly wage bill was reported for male and female employees but not months of operation, while in 1870 and 1880, the total annual wage bill and months of operation were reported. The change in definition occurs between 1860 and 1870, precisely when a jump in wage inequality occurs in the data.

Visual inspection of the 1850-60 wage distributions reveals a greater tendency towards heaping on integer values compared with the 1870-1880 distributions. Such heaping will bias the standard deviation downward relative to its true value, and thus potentially overstate the rise in wage inequality. We adjusted for the change in reporting by computing artificial distributions for 1870 and 1880 in which the establishment wage, as estimated, was rounded the nearest integer value. Compared with the actual distributions, the artificial distributions show greater heaping and are less dispersed as we expected. However, the artificial distributions still show greater wage dispersion when compared with either the 1850 or 1860 distributions. Changes in reporting procedures do not appear to be responsible for the rise in wage dispersion.

Of somewhat greater concern are sample selection biases. We can explore the ramifications of two such biases. First, certain industries in 1880 were supposed to be canvassed by "special agents" rather than by regular census enumerators. The agents were chosen for their knowledge of industry conditions, and it was thought (at the time) that such knowledge would improve the quality of enumeration. Unfortunately, the manuscript schedules collected by the agents have never been found, and are presumed lost (Delle Donne 1973). In practice, the distinction between "special agent" industries and other industries was not air-tight: some firms that should have been enumerated by special agents were, in fact, enumerated by regular Census employees (and appear in our 1880 sample). However, it is quite clear that certain industries are under-represented in the 1880 sample

and, therefore, it is possible that our results on wage inequality are affected by the under-representation. Fortunately, the identities of the "special agent" industries are known, so it is possible to exclude establishments in the relevant SIC codes.[6]

In Panel B of Table 1 we present inequality statistics for 1850 and 1880 excluding the special agent industries. Excluding these firms increases the level of inequality somewhat in 1850 relative to 1880, so that the rise in wage inequality is dampened. However, the increase is still impressive; for example, the 9-1 wage gap rose by 0.22 in log points (approximately 24.5 percent) if special agent industries are excluded, compared with 0.26 in log points (approximately 29.7 percent) if they are not.

Another possible bias concerns changes in the likelihood that very small establishments would be enumerated. The 1850 and 1860 censuses required that establishments produce at least $500 in output in order to be enumerated. Although the restriction was apparently intended to apply to the gross value of output (that is, including the value of raw materials) some enumerators may have interpreted it to apply to net value (value-added), thus failing to enumerate some small establishments in 1850 and 1860. The instructions to enumerators were clarified in 1870.

It is clear from inspecting the original schedules that not all enumerators were confused by the distinction between gross and net value of output because many small establishments were enumerated in 1850 and 1860. However, because the lower bound – $500 – was expressed in nominal terms, it is possible that changes in the price level (or in relative output prices across firms) did increased the likelihood of enumeration of very low-output establishments after 1850. If such firms paid wages that systematically differed from other firms, it is possible that the trends that we observe in wage dispersion might be biased. Unfortunately, we do not know exactly where "$500" was in the true distribution of output in 1850. In what we believe to be a conservative gesture, we assume that the cut-

off for such firms was at the 30th percentile of the distribution of gross output (that is, including the value of raw materials).[7] That is, for the purpose of comparison with the 1850 sample, we deleted establishments in the bottom 30 percent of the distribution of gross output in the 1880 sample, the assumption being that such firms would not have been enumerated in 1880 had the instructions to enumerators not been changed in 1870. The import of excluding such firms was to narrow the extent of wage dispersion slightly in 1880, implying a somewhat smaller increase over time in wage inequality. However, the effect was small: excluding low-output firms in 1880 results in an increase of 0.36 in log points in the 90-10 range statistic between 1850 and 1880, compared with approximately 0.39 if such are firms are kept in the 1880 sample (see Panel A). Although further work would certainly be useful, we conclude that the bulk of the rise in wage inequality shown in Table 1 is a real historical phenomenon, not just an artifact of changes in census enumeration procedures or sample selection biases.

3. Regression and Decomposition Analysis

We have documented that a rise in inequality in establishment wages occurred between 1850 and 1880. In order to account for the sources of the change in wage inequality we compute decompositions of the change over time in the range statistic based on regressions of the form

$$\{\text{eq. 1}\}: \ln w = X\beta + \varepsilon = X\beta + \sigma\delta$$

where the X's are a set of establishment level variables reported in the census, $\varepsilon = \sigma\delta$ is the residual,

σ is the standard error of the residual, and λ is the "standardized" residual ($\delta = \epsilon/\sigma$). Included in X are the percent female, the log of the number of workers and its square, the log of the capital-labor ratio, a dummy variable indicating the use of steam power, urban and regional dummies, and industry dummies measured at the 2-digit SIC level. Because Table 1 makes it clear that the increase in inequality in the 1860s persisted to 1880 we focus on the change between 1850 and 1880.

Following Juhn, Murphy, and Pierce (1993; see also Katz and Murphy 1992, and Goldin and Margo 1992), in interpreting {eq. 1} we think of β and σ as elements of the "wage structure" – that is, the coefficients attached to the observed (the X's) and unobserved (δ) characteristics; together, as defined by {eq. 1}, the two (coefficients and characteristics) determine ln w. Thus, inequality in ln w can be decomposed into components associated with the terms in {eq. 1} (see below).

In studies of the wage inequality based on individual-level data the independent variables included in wage regressions like {eq. 1} typically pertain to personal attributes of workers, such as schooling, labor market experience, race, gender, and so on, at least some of which (for example, schooling and experience) are meant to measure productive characteristics – that is, human capital. Except for the percent female, information on worker characteristics was not reported in the 19th century manufacturing censuses. However, recent work by Goldin and Katz (1998) suggests that several of the establishment-level variables in our regressions can be given plausible interpretations in terms of variations across establishments in the skill composition of workers. Goldin and Katz model the evolution of manufacturing as occurring in four stages -- artisan shop to factory to assembly line and, lastly, to "continuous process" production. While their model also incorporates factor substitution between capital and labor, the basic ideas are readily conveyed in a simpler three-stage version. There are three factors of production: capital, skilled ($L_s$), and unskilled ($L_u$) labor. In the

artisan stage, skilled labor used capital to fashion a final product.  In the factory stage, the process of

crafting the finished good was divided up into a series of steps, each of which could be performed by

an unskilled worker, possibly with the use of some machinery, either hand or water-powered; some

skilled labor was retained to maintain and repair any such machinery (and, possibly, produce it).  By

its very nature division of labor required more workers than the artisan shop; on average, factories

were larger (had more employees) than artisan shops. Because the factory was more productive than

the artisan shop, in switching from artisanal to factory production, the capital-output ratio may have

risen, but not necessarily the capital-labor ratio (Sokoloff 1984).   However, because the factory

process economized on the relative use of skilled labor,  $L_s/(L_s + L_u)$ -- the share of skilled labor --

was much lower in factories than in artisan shops.[8]  In the early years of American industrialization,

this substitution was accomplished primarily through the employment of native-born women (and

children), whose productivity compared with adult men was low in agriculture in the American

Northeast, where manufacturing first took hold (Goldin and Sokoloff 1982, 1984).  Beginning in the

late 1840s the labor supply of both groups was supplemented by immigrants, but the identification of

women and child labor with (low-skilled and low-wage) factory employment remained throughout

the century (Goldin 1990).

As manufacturing continued to evolve, some establishments adopted "continuous process"

technologies that required more capital per worker than factories, and also an inanimate power source.

 Holding constant the number of employees, continuous process establishments  used relatively more

skilled labor because of the need to maintain production "throughput" or interchangeable parts

machinery (Chandler 1977).  Goldin and Katz (1998) test their model using industry level data from

the early 20th century and find that the data are consistent with the model's predictions: skill intensity

was decreasing in firm size and positively related to electricity usage and the capital-labor ratio.

Although Goldin and Katz crafted their model to elucidate the underlying economic rationale for "capital-skill complementarity" in early 20th century manufacturing, they note that the evolutionary process just described began early in the 19th century; and that the hypothesized relationships between skill composition and establishment size, capital intensity, and energy use could well have applied prior to 1900. For example, steam power which diffused rapidly after 1850, used special machinery that required constant maintenance in the hands of skilled workers (see Hunter 1979).

Unfortunately, it is not possible to test the Goldin-Katz model for the data analyzed in this paper because we do not observe directly any measures of the skill composition of establishment workforces (other than, perhaps, the percent female). In an appendix, we provide some indirect evidence in favor of the model for 1880 by making use of the data on the daily wages of mechanics and common labor that were also reported in the census for that year.

Having estimated the regressions, we use the coefficients to compute a decomposition of the change in the 90-10 range statistic between 1850 and 1880. We first compute $\ln w^*$

$$\ln w^* = X_{1850}\beta_{1880} + \sigma_{1850}\delta_{1850}$$

The difference between the range statistics of $\ln w_{1850}$ and $\ln w^*$ reflects the impact of changes in the regression coefficients – the $\beta$'s – on the change in wage inequality between 1850 and 1880.

Next, we compute $\ln w^{**}$

$$\ln w^{**} = X_{1850}\beta_{1880} + \sigma_{1880}\delta_{1850}$$

15

The difference between the range statistics of ln w* and ln w** reflect any change over time $\sigma$. A rise in $\sigma$ indicates that "residual" wage inequality – that which cannot be explained by the X variables – increased over time; holding an establishment's position in the residual distribution ($\delta$) fixed, the distance in wage terms between that establishment and any other is now larger.

Finally, the difference in the range statistics of ln $w_{1880}$ and ln w** reflect changes over time in the distribution of establishment characteristics X, and the mapping between X and an establishment's position in the residual distribution. For example, if high-wage establishments were also capital-intensive establishments, and the distribution of establishments (employment-weighted, see below) shifted towards those that were highly capital-intensive, the range statistic would increase in value. Below we show that changes in establishment size – the log of the number of workers – functioned in this manner to produce a more unequal distribution of establishment wages in 1880 than in 1850.

3.1 Regression Results

We estimated regressions of the log of the establishment wage on the log of the number of employees and its square, the log of the capital-labor ratio, the percent female, and dummy variables for the use of steam power, urban status, region, and industry. Observations are weighted by the number of employees per establishment. Because the decomposition requires a consistently defined set of industry dummies in both years, the regressions are estimated using the samples excluding the 1880 "special agent" industries (see the notes to Panel B of Table 1).

The regression results are shown in Table 3. Both regressions do a moderately good job in

explaining the variation in wages across establishments, accounting for, respectively about 44 percent of the variance in 1850 and somewhat less – 33 percent – in 1880. The mirror image of the fall over time in $R^2$ was a rise in the standard error of the regression, from 0.34 in 1850 to 0.46 in 1880.

In preliminary regressions we included just a linear term in size (that is, the log of the number of workers). We found, in accordance with the Goldin-Katz model, a negative effect on the establishment wage in both years. However, an analysis of the residuals revealed that including size squared improved the fit of both regressions. In fact, as Table 2 shows, when size squared is included, the coefficient of size is positive, while that of size squared is negative. The inverted-U relationship between wages and size appears to have become more pronounced between 1850 and 1880 – the coefficients on both size and size squared as larger in absolute value in 1880 than in 1850.[9] In the Appendix we provide some evidence to suggest that the positive effect of the linear term (size) and the negative effect of the squared term (size squared) does not reflect, first, a rising share of skilled labor with size and then a decline; rather, the share of skilled labor declines continuously with size. Thus the positive coefficient must capture some positive effect of size on wages other than skill composition, an effect that may have been growing over time. Studies using 20th century data routinely find a positive correlation between firm size and wages, controlling for worker (and other establishment) characteristics (see, for example, Brown and Medoff 1989), so it is perhaps not too surprising that, over some range, wages and establishment size were positively correlated in 19th century manufacturing. However, the negative coefficient on size squared makes it clear that eventually wages declined with establishment size, in both 1850 and 1880.

Consistent with the Goldin-Katz model and much previous evidence (see, for example, Goldin and Sokoloff 1982), wages in female-intensive establishments were lower than male-intensive establishments: an increase of 10 percentage points in the percent female (in terms of the regression,

17

from, say, 0.5 to 0.6) lowered the establishment wage by slightly more than 4 percent.[10]  Also consistent with the model, capital intensive establishments paid higher wages, as did firms that utilized steam power.  The elasticity of the establishment wage with respect to capital intensity was about 0.10, indicating that a doubling of the capital-labor ratio was associated with an increase in the establishment wage of roughly 10 percent.  Controlling for other factors, steam-powered establishments paid wages that were 13 percent higher than non-steam powered establishments in 1850; by 1880, when steam power was more widely diffused, the earnings premium had declined to approximately 8 percent.

It is well-known that, in the immediate aftermath of the Civil War, factor markets were severely disrupted, and per capita incomes in the South declined relative to the North (Wright 1986; Bodenhorn and Rockoff 1992).  We expected, therefore, to find that wages at southern establishments declined relative to other regions between 1850 and 1880, especially for establishments in the South Atlantic states – the so-called "low-wage" South (see Wright 1986).  Instead, however, we found that, while wages were indeed lower than average in the South, wage gaps with respect to the North were either approximately the same in both years (the South Central states) or smaller in absolute value (the South Atlantic states).  According to the data examined in this paper, the South Atlantic states were already low-wage in 1850; they did not become more so over the period – indeed, wages in the South Atlantic converged on Northern levels, particularly the Northeast. Convergence was also true for establishments in the West, where wages were very high relative to other regions in 1850, but markedly less so in 1880.

An important feature of the changing geography of manufacturing in the 19th century  was a shift in production towards urban areas (Kim 2000).  The price of land was higher in urban locations; thus an urban location would have to confer a productivity advantage to offset higher non-labor costs.

Workers in 19th century firms tended to live close by where they were employed, and thus workers in urban firms would typically face a higher cost of housing than rural workers. Consequently, nominal wages – the data that we are analyzing – would need to be higher in urban firms.[11] The results for both regressions bear this out; there is some evidence that the urban-rural wage gap was larger in 1880 than in 1850, although the difference between the two years in this respect was small and arguably not economically significant (about 3 percent).

Notable features of 20th century manufacturing wages are the existence and stability of inter-industry wage differentials (see the references in Allen 1995). To investigate such differentials we included 18 industry dummies defined at the 2-digit SIC level; in the interests of space the coefficients are not reported.[12] Four features of the coefficients of the worth noting. First, differences across industries were important; F-tests indicate that coefficients of the dummy variables were jointly significant in both years. Second, the coefficients of the dummies were positively correlated across the two years ( r = 0.467, significant at the 6 percent level). Similar to modern data, high-wage industries remained high-wage -- conversely, for low-wage industries. Third, there is evidence of regression to the mean. Industries that were high-wage in 1850 were less so in 1860 (conversely for low-wage industries) although, as just noted, the auto-correlation was positive. Fourth, inter-industry differentials declined somewhat between 1850 and 1880; weighted by employment, the standard deviation of the industry coefficients of was 0.089 in 1880, compared with 0.102 in 1850.

3.2 Decompositions

We compute a decomposition of the change in the 90-10 range statistic between 1850 and 1880 using the method described earlier. The results are shown in Panel A of Table 3.

The first key finding is that changes in the regression coefficients, by themselves, produced a narrowing of the 90-10 range statistic -- that is, a reduction in wage inequality. It is apparent that the negative effect on inequality occurred primarily because of compression in the regional and industry coefficients. However, there is no evidence from the regressions that, in terms of the crucial variables in the Goldin-Katz model – size, capital intensity, use of steam power, the percent female – the "skill differential" increased. If this were true, we would expect to see uniformly larger coefficients (in absolute value) on all of these variables in 1880 than in 1850 – which is not the result that we obtain. However, this description of changes in the structure of establishment wages in terms of the regression coefficients is incomplete, because residual wage inequality increased. Overall, the two changes in the structure of wages largely offset each other, leaving the level of inequality of establishment wages only slightly higher in 1880 than in 1850.

By far the largest term in the decomposition is that pertaining to changes in the distribution of the independent variables; in log terms, fully 0.272 points of the 0.310 point increase in the 90-10 range statistic can be attributed to changes in the distribution of the independent variables. While the decomposition is well-suited to identifying overall "quantity" (the independent variables) versus "price" effects (the regression coefficients and the standard deviation) it is less well-suited to identifying which of the independent variables were primarily responsible for the rise in wage inequality.

This question can be most easily addressed by computing a different decomposition, that of the change in the 9-1 wage gap. Specifically, we compute the expression

$$\beta_{1880} [\Delta X_{1880} - \Delta X_{1850}]$$

where "Δ" refers to the difference in mean values between the ninth and first deciles of the relevant X variable for the given year. In this computation, the X variables are size and size squared, and the β's are their associated coefficients for 1880 – hence, the expression measures the change in the 9-1 wage gap that occurred because of changes in the distribution of establishment sizes. The results are shown in Panel B of Table 3..

The results in Panel B confirm that changes in the distribution of establishment sizes produced an increase in the 9-1 wage gap. High wage (that is, $9^{th}$ decile) establishments were more likely to be small than low wage (that is, 1rst decile) establishments in both years, but the gap in size between high and low wage establishments grew over time. Simply put, between 1850 and 1880 employment became more concentrated at the largest establishments.[13] The growing concentration of employment at large establishments produced an increase in wage inequality, as shown by its effect on the 9-1 wage gap -- that is, an increase in the average difference in wages between high and low-wage establishments.

4.0 Discussion

During the second half of the $19^{th}$ century the United States experienced substantial growth economic growth, in both aggregate and per capita terms. Labor productivity rose, and a pronounced shift of labor out of agriculture occurred. In particular, manufacturing's share of the labor force increased from approximately 14 percent in 1850 to 20 percent in 1900; indeed, fully 34 percent of all non-farm labor was employed in manufacturing at the start of the $20^{th}$ century.[14] Although data

21

problems make it is difficult to provide a precise estimate of the aggregate growth rate over any long interval, few (if any) economic historians would dispute the claim that, on average, real wages of manufacturing workers were substantially higher in 1880 than in 1820, when the American economy first began to industrialize (Williamson and Lindert 1980; Sokoloff and Villaflor 1992).

While it may seem obvious (at least to economic historians) that the average American worker in the 19th century was (in real wage terms) made "better off" by the growth of manufacturing, the impact of such developments on the wage structure is far from obvious. We find that the dispersion of establishment wages in manufacturing increased substantially between 1850 and 1880. The rise in wage inequality did not occur primarily because of changes in the relationships between the establishment wages and establishment characteristics such as capital intensity or the use of steam power which, we claimed (and provide some evidence for in the Appendix) were proxies for the skill composition of establishment workforces – that is, our results provide no evidence that the "skill differential" in manufacturing rose over the period.[15] However, as we noted earlier, movements in the skill differential – as captured by changes in the regression coefficients -- do not adequately summarize the structure of establishment wages over the 1850 to 1880 period, because residual wage inequality increased.[16] By its very nature it is difficult to explain why residual wage inequality rose. Certainly, it is possible that the rise in residual inequality represents a rise in the returns to skills embodied in workers not proxied by the independent variables in the regression.[17] In any event, it is clear from our results that, overall, changes in the wage structure do not account for the rise in establishment wage inequality between 1850 and 1880.

As manufacturing evolved after 1850, the share of employment in large establishments increased. Our decompositions demonstrate that the growing concentration of employment in large establishments was a major factor in the rise in inequality in the establishment wage. These were

precisely the establishments that were substituting operatives for skilled artisans, in the process of implementing production technologies that emphasized division of labor and which made increasing use of steam power and complex machinery. The former – division of labor – increased the share of the manufacturing workforce that was unskilled compared with the artisan shop, while the latter boosted the demand for skilled labor of various sorts – engineers, machinists, and a variety of other specialized technical occupations. Our results show that, in terms of the distribution of establishment wages, the compositional effects of size via division of labor – "de-skilling" – dominated, producing a rise in wage dispersion across establishments.

Our data on wage inequality, we reiterate, pertain to establishment wages, not to individual workers. The rise in wage inequality documented in this paper need not have translated into a rising inequality across individuals manufacturing workers, if wage inequality within establishments declined while wage inequality across establishments was on the rise. Unfortunately, we know of little hard evidence on this issue; nor is it likely that anything other than scattered archival sources could be found for the 1850 to 1880 period. However, while employment was becoming increasingly concentrated in large establishments, the median size establishment was still very small in 1880 (3 employees). Thus, if within-establishment inequality declined to a sufficient degree to offset the rise across establishments, the decline would have to have been concentrated in large establishments to be economically meaningful. But these were the establishments which had carried division-of-labor to its logical conclusion, subdividing tasks to a fine degree and reducing their dependence on skilled artisans. Because this process was an ongoing one it is not likely that wage differentials within establishments declined between 1850 and 1880; indeed, if anything it is more likely that such differentials were increasing.

It is always perilous to draw economy-wide inferences from sectoral data – doubly so, in our

case because much of the labor force in the United States between 1850 and 1880 was engaged in agriculture, either self-employed or otherwise not working for wages. Two points, however, can be made. First, workers who were self-employed in agriculture may have worked for wages in other sectors, such as manufacturing, during seasonal lulls in agricultural production. Certainly, there is little evidence of significant wage gaps between agriculture and other sectors, controlling for location and worker characteristics (Williamson and Lindert 1980). Second, for other periods in American history there is an impressive degree of co-movement between series on individual earnings inequality and the sort of data analyzed in this paper. Allen (1995) has shown that the time-series behavior of the coefficient of variation of average annual earnings across two-digit industries closely mimics measures of overall earnings inequality: good examples include the 1940s, when both the across-industry coefficient of variation and overall earnings inequality declined sharply, and the post-1970 period of rising wage inequality.

If the rise in wage inequality across manufacturing establishments is symptomatic of a more broadly based rise in wage inequality, the decomposition suggests that the fundamental source would have been an increase in the economy-wide proportion of workers who were "unskilled". Unfortunately, because of limitations of census evidence on occupations, it is not possible to produce, in a fully consistent manner, national estimates of the proportion of the labor force that was unskilled in 1850 compared with 1880. However, it is possible to produce estimates for adult white males. Table 5 shows estimates derived from the IPUMS (integrated public use microdata) samples of the 1850 and 1880 federal censuses, using two definitions of unskilled, one narrow and the other somewhat broader. These estimates are consistent with the hypothesis that, between 1850 and 1880, the share of the labor force that was unskilled increased.

After 1880 manufacturing plants grew ever larger in size, as continuous process production

24

methods started to diffuse in earnest (Nelson 1975; Chandler 1977). Later, electricity began to replace steam as the inanimate power source of choice. According to Goldin and Katz (1998), these changes boosted the relative demand for skilled labor and, by themselves, would have substantially increased the skill differential in manufacturing above the level prevailing at the end of the 19th century. However, the so-called "high school movement" intervened. Rapid growth in educational attainment associated with the movement greatly expanded the relative supply of educated labor. Many of the newly minted graduates entered the "glamour" manufacturing industries of the day, which had much higher skill requirements than old line industries of the first industrial revolution. However, the increase in relative supply was so great that the returns to skill were bid down and, accordingly, the 90-10 range in wages among manufacturing workers fell sharply after the turn of the century, as did the range of inequality in establishment wages (Williamson and Lindert 1980; Goldin and Katz 1999; Brissenden 1929).

5. Concluding Remarks

The 19th century is largely uncharted territory in the history of wage inequality in the United States.. Previous work has focused almost exclusively on movements in skill differentials, the ratio of skilled to unskilled wages (Williamson and Lindert, 1980; Margo 2000). By contrast, this paper has examined the dispersion in wages across manufacturing establishments between 1850 and 1880. We find that this dispersion increased substantially during this period, primarily because of an increasing share of employment in large establishments, which paid lower than average wages, rather than changes in the structure of wages across establishments. As such, this paper provides the first quantitative evidence that we are aware of that "de-skilling", as labor historians call it, altered the

distribution of wages in 19<sup>th</sup> century manufacturing.

This work could be extended in two directions. First, we have suggested that the trends in inequality in manufacturing uncovered in this paper may be representative of the economy as a whole. Although such a claim has a basis in parallel patterns found in later periods, as well as some evidence on changes over the period in the skill composition of the labor force, direct evidence for sectors other than manufacturing would clearly be useful. Second, we have emphasized the purely economic implications of our findings in this paper. However, the post-1880 period witnessed rapid growth in labor strife, culminating in rising rates of unionization. Large establishments were frequently the targets of strikes, shutdowns, and other forms of labor activism. The results of this paper suggest that this was no accident, because the growth of large establishments fueled growing wage inequality within manufacturing. But the precise links between the findings of this paper and the political economy of the labor movement in the late 19<sup>th</sup> century remain to be investigated.

6. Appendix

We have shown that a growing concentration of employment in large establishments, which paid lower wages on average, can explain the rise in wage inequality across establishments between 1850 and 1880. Following Goldin and Katz (1998), we have interpreted this finding in terms of differences in the skill composition of workforces between large and small establishments – because of division of labor, large establishments, other factors held constant, used a higher proportion of unskilled labor than small establishments. While just such an effect has long been pointed out by labor historians – indeed, it is precisely what is meant by "deskilling" – quantitative evidence from the period is lacking, because none of the extant 19<sup>th</sup> century manufacturing censuses collected direct

distribution of wages in 19th century manufacturing.

This work could be extended in two directions. First, we have suggested that the trends in inequality in manufacturing uncovered in this paper may be representative of the economy as a whole. Although such a claim has a basis in parallel patterns found in later periods, as well as some evidence on changes over the period in the skill composition of the labor force, direct evidence for sectors other than manufacturing would clearly be useful. Second, we have emphasized the purely economic implications of our findings in this paper. However, the post-1880 period witnessed rapid growth in labor strife, culminating in rising rates of unionization. Large establishments were frequently the targets of strikes, shutdowns, and other forms of labor activism. The results of this paper suggest that this was no accident, because the growth of large establishments fueled growing wage inequality within manufacturing. But the precise links between the findings of this paper and the political economy of the labor movement in the late 19th century remain to be investigated.

6. Appendix

We have shown that a growing concentration of employment in large establishments, which paid lower wages on average, can explain the rise in wage inequality across establishments between 1850 and 1880. Following Goldin and Katz (1998), we have interpreted this finding in terms of differences in the skill composition of workforces between large and small establishments – because of division of labor, large establishments, other factors held constant, used a higher proportion of unskilled labor than small establishments. While just such an effect has long been pointed out by labor historians – indeed, it is precisely what is meant by "deskilling" – quantitative evidence from the period is lacking, because none of the extant 19th century manufacturing censuses collected direct

evidence on the skills of manufacturing workers.

However, it is possible to make use of some additional information reported in 1880 to estimate, albeit indirectly, the determinants of skill composition across establishments. As noted earlier in the paper, the 1880 census also reported the average daily wage of "common labor" and the average daily wage of "mechanics". Let $\lambda$ = share of "skilled" workers in establishment j, and $w_{ij}$ be the average daily wage, i = s(killed) and u(nskilled). We have the following identity:

$$\{eq.\ 2\}:\ w_j = D_j(w_{sj}\lambda + (1-\lambda)w_{uj})$$

where, as before, $w_j$ is the establishment (monthly) wage, and $D_j$ = average number of days in operation per month. Replacing $w_{sj}$ by the daily wage of mechanics, $w_{uj}$ by the daily wage of common labor, $D_j$ by a constant $\underline{D}$, and letting $\lambda = X\beta$, we have, after re-arranging terms, the following regression

$$\{eq.\ 3\}:\ w_j = \alpha(w_c + (w_m - w_c)X\beta)$$

where c = common labor and m = mechanics. This wage regression differs from the 1880 regression reported in Table 2 in three ways: first, the dependent variable is in levels, not logs; second, there is no intercept; and third, the two daily wages are included as right-hand side variables, as specified.

Table 5 shows the results of estimating {eq. 3} in which the X variables are the same as those included in the wage regressions in Table 2. Caution should be exercised in interpreting the results, in particular, because the average daily wage of mechanics is undoubtedly a less than perfect proxy

for the average daily wage of all "skilled" employees. Caveats aside, the results are supportive of the Goldin-Katz model. Larger establishments were less skill intensive: note, in particular, that the coefficient on the linear term (the log of the number of workers) is negative, as is the coefficient of size squared. The fact that the coefficient of the linear term is negative in Table 5 suggests that the positive coefficient of the same variable in Table 2 captures features of establishment size that were unrelated to skill composition – for example, the work environment – whose influence on wages apparently intensified between 1850 and 1880. We also find that establishments that employed female labor were less skill-intensive, while capital-intensity and use of steam power were positively related to skill intensity. Little evidence is found of significant differences in skill intensity between regions in the North, or between rural and urban establishments; however, there is some evidence of regional differences in the South, with establishments in the South Atlantic states choosing to employ a higher proportion of unskilled labor. In the interests of space, we do not report the coefficients of the industry dummies; however, the correlation between these and the coefficients of the industry dummies from the log wage regressions reported earlier was positive, indicating that some of the inter-industry variation in wages was due to differences in skill composition.[18]

References

Allen, Steven G. 1995. "Updated Notes on the Inter-industry Wage Structure, 1890-1990," Industrial
    and Labor Relations Review 48 (January): 305-320.

Atack, Jeremy and Fred Bateman. 1999. "U.S. Historical Statistics: Nineteenth Century U.S. Industrial
    Development Through the Eyes of the Census of Manufactures," Historical Methods 32 (Fall):
    177-188.

Bodenhorn, Howard and Hugh Rockoff. 1992. "Regional Interest Rates in Antebellum America." In
    C. Goldin and H. Rockoff, eds. Strategic Factors in Nineteenth Century American Economic
    History: A Volume to Honor Robert W. Fogel, pp. 159-187. Chicago: University of Chicago
    Press.

Brissenden, Paul F. 1929. Earnings of Factory Workers, 1899-1927: An Analysis of Pay-Roll
    Statistics. Washington: Government Printing Office.

Brown, Charles and James Medoff. 1989. "The Employer Size-Wage Effect," Journal of Political
    Economy 97 (October): 1027-1059.

Brown, Martin and Peter Phillips. 1986. "Craft Labor and Mechanization in Nineteenth-Century
    American Canning," Journal of Economic History 46 (September): 743-756.

Chandler, Alfred. 1977. The Visible Hand: The Managerial Revolution in American Business.
    Cambridge: Harvard University Press.

Coehlo, Phillip and James Shepherd. 1976. "Regional Differences in Real Wages: The United States,
    1851-1880". Explorations in Economic History 13 (April): 203-30.


Engerman, Stanley and Claudia Goldin. 1993. "Seasonality in Nineteenth Century Labor Markets,"

in D.Schaefer and T. Weiss, eds. <u>Economic Development in Historical Perspective</u>. Stanford, CA: Stanford University Press.

Gerber, James. 1997. "Agricultural Expansion During the Gold Rush: California Grain Farming as a 'Booming' Lagging Sector." Unpublished paper, Department of Economics, San Diego State University.

Goldin, Claudia. 1990. <u>Understanding the Gender Gap: An Economic History of American</u> <u>Women</u>. New York: Oxford University Press.

Goldin, Claudia and Lawrence Katz. 1998. "The Origins of Technology-Skill Complementarity," <u>Quarterly Journal of Economics</u> 113 (June): 693-732.

Goldin, Claudia and Lawrence Katz. 1999. "The Returns to Skill in the United States Across the Twentieth Century," National Bureau of Economic Research Working Paper No. 7126, May, Cambridge, MA.

Goldin, Claudia and Robert A. Margo, 1992. "The Great Compression: The Wage Structure in the United States at Mid-Century," <u>Quarterly Journal of Economics</u> 107 (February): 1-34.

Goldin, Claudia and Kenneth Sokoloff. 1982. "Women, Children, and Industrialization in the Early Republic: Evidence from the Manufacturing Censuses," <u>Journal of Economic</u> <u>History</u> 42 (December): 741-774.

Goldin, Claudia and Kenneth Sokoloff. 1984. "The Relative Productivity Hypothesis of Industrialization," <u>Quarterly Journal of Economics</u> 99 (August): 461-88.

Habbakuk, H.J. 1962. <u>American and British Technology in the Nineteenth Century</u>. Cambridge: Cambridge University Press.

Hunter, Louis C. 1979. <u>A History of Industrial Power in the United States, 1780-1930. Volumes</u> <u>1 and</u> <u>2</u>. Charlottesville, Virginia: University Press of Virginia.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. "Wage Inequality and the Rise in the Returns to Skill," Journal of Political Economy 101 (June): 385-409.

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

Kim, Sukkoo. 2000. "Urban Development in the United States, 1690-1990." Southern Economic Journal 66 (April): 855-880.

Lebergott, Stanley. 1964. Manpower in Economic Growth: The American Record Since 1800. New York: McGraw-Hill.

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

Nelson, Daniel. 1975. Managers and Workers: Origins of the New Factory System in the United States, 1880-1920. Madison, WS: The University of Wisconsin Press.

Rosenbloom, Joshua. 1996. "Was There a National Labor Market at the End of the Nineteenth Century? New Evidence on Earnings in Manufacturing," Journal of Economic History 56 (September): 626-656.

Sokoloff, Kenneth. 1984. "Was the Transition from the Artisanal Shop to the Nonmechanized Factory Associated with Gains in Efficiency? Evidence from the U.S. Manufacturing Censuses of 1820 and 1850," Explorations in Economic History 21 (October): 329-350.

Sokoloff, Kenneth and Georgia C. Villaflor. 1992. "The Market for Manufacturing Workers During Early Industrialization: The American Northeast, 1820 to 1860." In C. Goldin and H. Rockoff, eds. Strategic Factors in Nineteenth Century American Economic History: A Volume to Honor Robert W. Fogel, pp. 29-65. Chicago: University of Chicago Press.

U.S. Department of Commerce. 1975. Historical Statistics of the United States. Washington: Bureau

of the Census.

Williamson, Jeffrey G. and Peter H. Lindert. 1980. <u>American Inequality: A Macroeconomic</u> <u>History</u>.

New York: Academic Press.

Wright, Gavin. 1986. <u>Old South, New South.</u> New York: Basic Books.

Table 1: Wage Inequality Across Manufacturing Establishments, 1850 and 1880: Log of Monthly Establishment Wage

| Panel A: Establishments with at least one worker | | | | |
|---|---|---|---|---|
|  | 1850 | 1860 | 1870 | 1880 |
| 90-10 range statistic | 1.13 | 1.168 | 1.307 | 1.514 |
| 9-1 wage gap | 1.649 | 1.676 | 1.906 | 1.909 |
| σ | 0.478 | 0.481 | 0.549 | 0.557 |
| N(establishments) | 5,481 | 5,313 | 3,731 | 6,519 |
| N(workers) | 44,128 | 48,746 | 43,862 | 78,219 |
| Panel B. Excluding special agent industries | | | | |
| 90-10 range statistic | 1.204 |  |  | 1.514 |
| 9-1 wage gap | 1.691 |  |  | 1.91 |
| σ | 0.491 |  |  | 0.557 |
| N(establishments) | 5,106 |  |  | 6,350 |
| N(workers) | 32,705 |  |  | 77,758 |

Notes:

Source: Atack and Bateman (1999) samples of manuscript censuses of manufacturing; see text.

90-10 range statistic: difference between 90th and 10th percentiles, log of average monthly wage

9-1 wage gap: difference in mean log wage, 9th decile minus 1rst decile

σ: standard deviation

Observations are weighted by employment.

In both panels, certain very low-wage firms are excluded (the cut-offs are 1850, $3.65; 1860, $3.98; 1870, $8.21; 1880, $8.07) and very high-wage firms are top-coded (the top-codes are 1850, $100; 1860, $125; 1870, $175; 1880, $166.67). Top-coded wages are replaced with values equal to 1.4 times the top-code (see Goldin and Margo (1992) for a similar procedure). In Panel B: firms in so-called "special agent" industries are excluded; see text.

Table 2: Wage Regressions: 1850 and 1880

| | 1850 | | 1880 | |
| --- | --- | --- | --- | --- |
| | $\beta$ | t-statistic | $\beta$ | t-statistic |
| Constant | 2.174 | 43.918 | 2.387 | 22.070 |
| Size | 0.059 | 4.661 | 0.159 | 12.159 |
| $Size^2 x\ 10^{-1}$ | -0.147 | -7.875 | -0.240 | -14.929 |
| % women | -0.527 | -17.842 | -0.440 | -14.218 |
| ln (capital/labor) | 0.106 | 20.532 | 0.090 | 15.315 |
| Steam power? | 0.142 | 9.023 | 0.077 | 5.357 |
| Urban | 0.181 | 10.253 | 0.210 | 14.197 |
| Midwest | -0.036 | -2.444 | 0.054 | 3.776 |
| South Atlantic | -0.329 | -18.407 | -0.197 | -8.680 |
| South Central | -0.093 | -4.027 | -0.110 | -3.070 |
| West | 1.414 | 22.107 | 0.244 | 6.851 |
| N | 5,036 | | 6,346 | |
| Adjusted $R^2$ | 0.514 | | 0.325 | |
| $\sigma$ | 0.342 | | 0.458 | |

Dependent variable is the natural logarithm of the monthly establishment wage; see text. Observations are weighted by employment. Size= log (employment). Regressions include 18 industry dummies coded at the 2-digit SIC level (see Atack and Bateman 1999 and footnote x). N: number of establishments.

Table 3: Decomposition of Change in Wage Inequality, 1850-1880

A. The Change in the 90-10 Range Statistic: Log of the Establishment Wage

|  | Value |
| --- | --- |
| Total Change, 1850 to 1880 | 0.310 |
| Due to regression coefficients | -0.118 |
| Due to standard deviation | 0.156 |
| Remainder (independent variables and standardized residuals) | 0.272 |

B. The Role of Establishment Size: Explaining the Rise in the 9-1 Wage Gap

|  | |
| --- | --- |
| Total Change, 1850 to 1880 | 0.219 |
| Change in 9-1 gap in: | |
| Size | -1.559 |
| Size squared | -17.572 |
| $\beta_{1880} (\Delta X_{1880} - \Delta X_{1850})$ | 0.177 |

Notes:

Panel A: see text for description of terms in the decomposition

Panel B: $\beta_{1880} (\Delta X_{1880} - \Delta X_{1850}) = 0.159 \times (-1.559) - 0.0240 \times (-17.572) = 0.177$. Change in 9-1 gap in size and size squared: the change between 1850 and 1880 in the difference in the mean values of size and size squared of establishments in the $9^{th}$ and 1rst deciles of the distribution of establishment wages.

Table 4: Aggregate Estimates of the Proportion Unskilled: White Males, Ages 20-60, in 1850 and 1880

| Definition of "Unskilled" | 1850 | 1880 |
|---|---|---|
| Narrow | 0.142 | 0.194 |
| Broad | 0.163 | 0.244 |

Source: IPUMS samples for 1850 and 1880. The narrow definition of unskilled includes only individuals who reported occupations of agricultural laborer, laborer (not elsewhere classified), or operative in manufacturing establishment (not elsewhere classified). The broader definition includes herders, teamsters, janitors, and individuals enumerated as manufacturing operatives in a specific industry (a list is available from Robert Margo on request).

Table 5: Skill Composition Regression: 1880

| Variable | Coefficient | T-statistic | Scaled Coefficient |
|---|---|---|---|
| $w_c$ | 21.628 | 7.880 | |
| $w_d = w_m - w_c$ | 7.141 | 1.732 | 0.330 |
| $w_d$ x Size | -1.119 | -2.584 | -0.052 |
| $w_d$ x Size$^2$ | -0.199 | -3.761 | -0.009 |
| $w_d$ x % women | -5.264 | -5.079 | -0.243 |
| $w_d$ x Log (K/L) | 0.501 | 2.528 | 0.023 |
| $w_d$ x Steam Power | 2.918 | 6.622 | 0.135 |
| $w_d$ x Urban | 2.998 | 5.196 | 0.139 |
| $w_d$ x Midwest | -0.049 | -0.103 | -0.002 |
| $w_d$ x South Atlantic | -2.646 | -3.088 | 0.122 |
| $w_d$ x South Central | 1.401 | 1.468 | 0.065 |
| $w_d$ x West | 0.603 | 0.615 | 0.028 |
| N (establishments) | 6,346 | | |
| $R^2$ | 0.775 | | |

See Appendix. Dependent variable is the monthly establishment wage; $w_c$ = daily wage of common labor, $w_m$ = daily wage of mechanics.  Scaled: coefficient/21.628.  Observations are weighted by employment.

1. By "structure of wages" we mean a set of prices associated with observed (and possibly unobserved) characteristics of individuals that determine earnings (see section 3). Wage "inequality" refers to differences in wages across some unit of observation (in this paper, manufacturing establishments); such differences depend on differences in characteristics and on the structure of wages.

2. In particular, the major sources of series of skill differentials are the so-called "Weeks" report of the 1880 census and the "Aldrich" report of 1893. Both are retrospective compilations of surviving payrolls of manufacturing establishments that were in existence at the time of either survey. Neither is a random sample, and both surveys substantially under-represent establishments outside the Northeast. For discussion of both surveys, see Lebergott (1966) and Margo (2000).

3. By top-coding we mean that observations in excess of the top-code are replaced by a multiple of the top-code (in our case, 1.4 times the top-code value; see Goldin and Margo (1992).

4. The need to remove outliers was also noted by Census Bureau in its 1929 study: "[t]hese extreme items are not only unrepresentative cases; they are to a certain extent spurious cases, representing perhaps establishments having only one or two employees, and in operation for, possibly, only a week or two during the year" (Brissendon 1929, p. 40).

5. To see this point, for example, in the case of the variance, let $w_{ij}$ = wage of the ith worker in the jth establishment, $n_j$ = number of workers in the jth establishment, and K = number of establishments. The variance of $w_{ij}$, is

$$\text{Var}(w_{ij}) = \sum_{j}^{K} \sum_{i}^{n_j} (w_{ij} - w_{..})^2/N$$

where $N = \Sigma\, n_j$, the number of workers, and the "." notation indicates the sample mean taken over the i or j subscripts – or both simultaneously, the grand mean, the weighted average of the establishment-level means. The variance can be decomposed:

$$\text{Var}(w_{ij}) = \text{within establishments} + \Sigma\, n_j(w_{.j} - w_{..})^2/N$$

It is the second term, the weighted variance across establishments, that can be studied with the data at hand.

6. The three-digit SIC codes of the "special agent" industries are silk and cottons (SIC 221, 222, 223, 225, 227, 228, and 229), coke (SIC 492 and 331), glass (SIC 321), ship-building (SIC 373), and distilleries and breweries (SIC 208).

7. $500 is approximately the 25th percentile of the distribution of output in 1850 – hence the claim in the text that establishing a cut-off at the 30th percentile is conservative.

8. In the Goldin and Katz model, the assembly line is the logical end of the "pure" factory, that is, economies of scale arising solely from division of labor, possibly aided by machinery.

9. However, if just the linear term is included, the coefficients are both negative (as noted above) and the 1850 coefficient is larger in absolute value ($\beta$ = -0.0356) than the 1880 coefficient ($\beta$ = -0.0248).

10. The magnitudes of the coefficients imply that, holding other factors constant, an all-female workforce would earn roughly 62-65 cents for every dollar earned by an all-male workforce. This figure is broadly consistent with previous work; see, for example, Goldin and Sokoloff (1982). In interpreting the coefficients it is important to keep in mind that the census did not consistently enumerate children separately over time; thus, we are only able to control for gender, not the age structure of the workforce.

11. An urban wage advantage could also arise if the average worker at urban establishments was more skilled than at rural establishments; see, however, section 3.3, where we present evidence for 1880 that differences in skill intensity between urban and rural establishments were statistically insignificant.

12. An appendix containing the coefficients is available on request from Robert Margo.

13. For example, in 1880, 41.3 percent of workers in the sample worked at establishments employing 100 or more workers, compared with 30.8 percent in 1850.

14. Computed from Lebergott (1966, p. 510). Recent revisions to Lebergott's estimates by Weiss (1992) would alter these figures somewhat but would not reverse the basic trend.

15. Because our 1880 sample reports the average daily wage of mechanics and common laborers, we can also estimate the skill premium (mechanics relative to common labor) in manufacturing in 1880. Our estimate is 1.73, very close to Williamson and Lindert's (1980) However, the data used by Williamson and Lindert give inadequate coverage to establishments located outside the Northeast before the Civil War. Margo (2000, ch. 3) has recently produced regional estimates of daily wages of common labor and skilled artisans for the 1850 census year; while these do not pertain to manufacturing per se, their geographic coverage is far superior to the data used by Williamson and Lindert. If Margo's regional estimates of the skill differential (artisan-to-common labor) are weighted by the regional shares of manufacturing employment found in our 1850 sample, the estimate of the aggregate skill differential is 1.68, slightly (3 percent) below our estimate for 1880. For this calculation we formed regional estimates of the skill differential using Margo's

wage estimates for 1849 (which, in principle, correspond to the 1850 census year); see Margo (2000, Tables 3A.5-6). In computing the regional employment shares for the 1850 census year, we exclude the very small proportion of establishments located in the far West.

16. Our finding that the "between" component of establishment wages (the regression coefficients) declined between 1850 and 1880 while the "within" component (residual inequality) is reminiscent of Juhn, Murphy, and Pierce's (1993) very similar result for the 1970s, which also witnessed declines in between-group wage inequality and rises in within-group wage inequality. Juhn, Murphy, and Pierce interpret their result as evidence that a uni-dimensional (or "single index") model of skill cannot adequately describe wage variation in the 1970s.

17. It is also possible (and difficult to refute) that, for example, the marginal impact of capital intensity on skill composition was smaller in 1880 than in 1850, but the skill differential was higher, such that the marginal impact of capital intensity on wages was the same. Our wage regression would not register any change in the capital intensity coefficient; however, to the extent that the residual captures unmeasured variation in skill composition, the rise in the skill differential would show up as a rise in the standard deviation, that is, in residual wage inequality.

18. The coefficient of the unskilled wage is an estimate of the average number of days per month in operation, approximately 22 days per month. Year-round operation in manufacturing during the period is generally assumed to have been between 300 and 310 days per year, or between 25 and 26 days per month. Thus the regression suggests that part-time operations generally took the form of fewer months per year – seasonality – rather than substantially fewer days per month; see Engerman and Goldin (1993).