NBER WORKING PAPER SERIES

DE-SKILLING: EVIDENCE FROM LATE NINETEENTH CENTURY AMERICAN MANUFACTURING

Jeremy Atack
Robert A. Margo
Paul Rhode

Working Paper 31334
http://www.nber.org/papers/w31334

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2023

We are grateful to workshop participants at Harvard University, the Federal Reserve Bank of Cleveland, and the University of Groningen for helpful comments on earlier versions of this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

De-skilling: Evidence from Late Nineteenth Century American Manufacturing
Jeremy Atack, Robert A. Margo, and Paul Rhode
NBER Working Paper No. 31334
June 2023
JEL No. J23,N61,O31

ABSTRACT

The long-standing view in US economic history is the shift in manufacturing in the nineteenth century from the artisan shop to the mechanized factory led to “labor deskilling.” Craft workers were displaced by mix of semi-skilled operatives, unskilled workers, and a reduced force of mechanics to maintain the powered machines. Investigating the Department of Labor’s 1899 Hand and Machine Labor Study using causal inference statistical techniques, we show the adoption of inanimate power did indeed induce deskilling. While the effects were statistically significant, they accounted for only 7-15 percent of the deskilling observed in the sample. Broadening the scope of our inquiry, we find the increased division of labor as captured by the increase in scale of operations and the ratio of workers to tasks accounts for a larger fraction.

Jeremy Atack
Vanderbilt University
6504 Hidden Hollow Trail
Brentwood, TN 37027
United States
and NBER
jeremy.atack@vanderbilt.edu

Robert A. Margo
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
margora@bu.edu

Paul Rhode
Economics Department
University of Michigan
205 Lorch Hall
611 Tappan St.
Ann Arbor, MI 48109-1220
and NBER
pwrhode@umich.edu

The traditional view among labor and economic historians is that the rise of the steam-powered factory encouraged "deskilling". That is, labor previously supplied by a skilled artisan using simple hand tools was displaced by special-purpose machinery powered inanimately and tended by a semi-skilled operative while unskilled workers performed various complementary tasks, such as shoveling coal to "feed" steam engines or moving partially finished goods or raw materials from one part of the shop floor to another. To be sure, the construction and installation of special-purpose machinery and its maintenance and repair required skilled mechanics but the substitution of semi- and unskilled dominated and the demand for skilled artisans relative to semi-skilled operatives and unskilled labor fell with the rise of the mechanized factory (Goldin and Katz 1998).

According to Andrew Ure, an early commentator,

> "[t]he principle of the factory system … is, to substitute mechanical science for hand skill, and the partition of a process into its essential constituents, for the division or graduation of [labor] among artisans. On the handicraft plan … skilled [labor] was usually the most expensive element of production.... but on the automatic plan, skilled [labor] gets progressively superseded, and will, eventually, be replaced by mere overlookers of machines" (Ure 1835, p. 20).[1]

[1] After 1890 electricity began to displace steam, and capital and skilled labor became relative complements as electrical power allowed factory owners to dispense with large numbers of unskilled workers (Goldin and Katz 1998).

More recently, Katz and Margo (2014) use occupational information from the US population censuses to investigate aggregate de-skilling in American manufacturing during the second half of the nineteenth century.[2] They further classify manufacturing workers into four categories – white collar, skilled blue collar, semi-skilled operatives, and unskilled laborers. Between 1850 and 1910, the skilled blue-collar share declined by 17 percentage points and the share of operatives and common labor increased by 7 percentage points – clear evidence of production de-skilling in the traditional sense (see Katz and Margo, 2014, Table 1.4).[3]

Data from the federal censuses of manufacturing show that, over the second half of the nineteenth century the share of establishments using steam power increased substantially (Atack, Bateman et al. 2008), suggesting that trends in de-skilling and mechanization were coincident. However, the establishment-level data from the censuses of manufacturing during the second

[2] The Census of Population began collecting occupational data in 1850 but did not specifically collect information on the industry of employment until 1910 (although the number of persons in household working in agriculture, commerce, or manufactures was reported in 1820). However, the handwritten occupation titles given on the original enumeration forms are sufficiently detailed that it is possible to classify workers into the manufacturing sector with a high degree of accuracy; see Katz and Margo (2014).

[3] The decrease in the skilled blue-collar share was not offset fully by the rise in the semi-skilled/operative share because the white-collar share also increased. Katz and Margo attribute the relative increase in white collar workers to the rise in scale – larger establishments were more complex to manage, requiring much more record-keeping and scheduling (hence, more white-collar workers) than did the artisan shop.

half of the nineteenth century provide no direct information on worker occupations or production tasks.[4] Nor is it possible to link the manufacturing returns directly to the population census, because the latter did not ask specifically where individuals worked. Hence, census data *per se* cannot address whether mechanized factories used relatively less skilled labor overall, and specifically for those production operations that were inanimately powered.

To investigate the role, if any, that mechanization played in de-skilling, we turn to an entirely different source, the U.S. Commissioner of Labor's 13th Annual Report (U.S. Department of Labor. 1899) on *Hand and Machine Labor* (hereafter HML study). We have recently digitized these data and made them amenable to quantitative analysis (Atack, Margo et al. 2023). This report was ordered by the U.S. Congress in 1894 to "investigate and report upon the effect of the use of machinery upon … the relative productive power of hand and machine labor" (U.S. Congress 1894). To this end, the Department's agents collected detailed data on the manufacture of several hundred highly specific manufactured products using the "machine methods" of most modern mechanized factory at the individual task level, recording the production operations from start to finish (U.S. Congress 1894, 1: 11). They then collected the same kind of information for the exact same product (or as close as possible) from establishments using the hand methods then "going out of use" (while remarking upon "the extent of the hand method of production, even at the present time") (U.S. Congress 1894, 1: 6).

[4] The Census of Manufactures began collecting data on inanimate power use in 1850 along with some information on worker age and gender. However, the available information is too coarse to directly address the de-skilling hypothesis, and there are other reasons why gender and age might be correlated with inanimate power use (see Katz and Margo 2014).

From these data, the HML staff generated a crosswalk linking the operations across the two production methods ordered from the tasks in the machine method of production to hand. This allows us to compare the occupations (as well as worker characteristics like age and gender) of those performing the equivalent tasks in traditional hand manufacture with those under machine production.

For each task, the HML staff recorded the occupation(s) of those performing it. These ranged from general titles such as "blacksmith" and "laborer" to particular occupations such as "drop forger" and "shank polisher." We have used these occupational titles to map the activity into codes based on those used by the 1950 Census, a coding scheme that traces its lineage back to the work of Alba Edwards for the 1910 Census who devised it as a means of classifying workers according the skill levels required by their jobs (United States. Bureau of the Census., Hunt et al. 1915; Edwards 1917). Using our codes, we allocate occupation titles into one of four bins – non-professional white collar, skilled artisans, semi-skilled operatives; and unskilled laborer. Our procedures are described in greater detail in Appendix A.

Based upon these allocations, we compute the share of production time performed by workers with these different skill levels separately for hand and machine production. For this purpose, we focus on the set of operations that overlapped between hand and machine production, for a sample size of 4,405 operations. We create a variable, *De-Skill* that equals one if the share of time performed by semi-skilled operatives and common labor in the operation exceeds the share so performed under hand labor. Our first finding is that the mean value of *De-Skilling* is 0.356, implying that in 36 percent of the matched operations, less skilled labor time was used in machine production than in hand production – clear evidence that de-skilling accompanied the transition from the artisan shop to the mechanized factory, as the traditional account maintains.

Our main analysis commences with ordinary least squares (OLS) regressions of *De-Skill* with fixed effects for the specific individual goods (called "units" in the HML study).[5] We find that (1) the more frequent use of inanimate power under machine labor is positively and significantly associated with higher de-skilling but that (2) this effect accounted for just 7 percent of the mean value (0.355) of *De-Skill*. Further probing using an instrumental variables estimator suggests that the OLS effect of mechanization is biased downwards compared with the true causal impact, but the "percent explained" of the mean value of *De-Skill* using our preferred IV coefficient is relatively modest -- about 15 percent. In short, while mechanization played a causal role in de-skilling, that role was quantitatively small – a surprising finding, given the importance attributed to mechanization in the traditional historical accounts.

To explore additional factors behind de-skilling we follow the procedure adopted in Atack, Margo, and Rhode (2022) in modifying our base regression to include additional variables, among which are measures of the division of labor. It is natural to consider division of labor as a factor behind de-skilling because the very essence of division of labor is dividing up operation tasks among separate workers rather than having them performed by a single more skilled artisan. Conceptually, as a factor behind de-skilling, the division of labor is distinct from mechanization, because it could occur independently, for example, in response to transportation improvements (Atack, Haines et al. 2011). There is sufficient information in the HML study to

[5] The empirical analyses in this paper complement our study of labor productivity differences between hand and machine labor in the HML study data; see Atack, Margo, and Rhode (2022).

construct measures of the division of labor at the unit level and so we include these in our modified regression specification. These show that increases in the division of labor are strongly associated with higher levels of de-skilling. Overall, we find that the additional factors, particularly the division of labor, to be far more important than mechanization *per se* in explaining de-skilling. This conclusion is important, because the trend towards greater division of labor was clearly underway well before the widespread diffusion of inanimate power in the decades following the Civil War (Fenichel 1966; Atack, Bateman et al. 1980).

The Hand and Machine Labor Study: A Brief Introduction

In this section we present a brief introduction to the *Hand and Machine Labor* study, along with a discussion of the information on occupations. Further details on the study can be found in our prior work (Atack, Margo et al. 2019; Atack, Margo et al. 2022; Atack, Margo et al. 2023).

Published in two volumes totaling approximately 1,600 pages, the HML study described the operations, or tasks involved in the production of what the study called “units”. These units were quantities of highly specific products such as “50 dozen regular taper, triangular saw files, 4 inches long, tapering 23/64 inch” (U.S. Department of Labor. 1899, 1: 241-6 and 2: 1026-9). For each unit, production was standardized between the two modes to industry norms by making adjustments solely to the aggregate time. The report covered 672 units in various economic sectors, 626 (units 28–653) of which were in manufacturing and are our focus here. Although the products in the study cover a wide array of industries (2-digit SIC codes 20–39) they are not a representative sample of manufacturing output at the time (see Atack, Margo et al. 2022, online Appendix, Table 1). The core of these manufacturing data is contained in a series of tables stretching across both the left and righthand pages of Volume Two, covering over 1,100 pages and set in small type in 13 columns (some of which contain multiple entries). An example of part of one of these tables is given in Figure 1.

For each unit, the HML staff actually collected production data from four establishments, two of which used "hand labor" methods and two "machine labor" methods.[6] To guide the collection efforts, the 1894 HML survey instructions noted: "Roughly characterized <u>hand labor</u> was the system of the last century and <u>machine labor</u> is the system of this…" (United States. Commissioner of Labor. 1894, Instruction form, DL 405, BLS Scrapbook, emphasis added). Further, the survey instructions stipulated "The machine schedule should be made for factories where the most highly developed machinery is in use, and the hand schedule should be for establishments using the most old fashioned and primitive of hand methods." (United States. Commissioner of Labor. 1894, Supplemental Instruction No. 2, DL 416, BLS Scrapbook).[7]

Almost all data on machine labor pertained to the 1890s. For hand labor, however, the staff relied primarily on historical data prior to the 1890s, one unit as early as 1813 although about a quarter were approximately contemporaneous with the machine methods.[8]

---

[6] The staff chose "the better and more complete" accounting of each mode of production for the published report (U.S. Department of Labor. 1899, 1: 13).

[7] Or as Wright put it, between the "most modern machine method[s]" compared with the "old fashioned hand process … in vogue before the general use of automatic or power machines" (U.S. Department of Labor. 1899, 1: 11).

[8] A small number of observations (15) on hand labor used data from foreign countries. These are excluded from our analysis.

The principal goal of the HML study was to measure differences in production times for specific operations, and overall, for each unit, between machine and hand labor. The data on production times collected by the staff demonstrated that machine labor was, on average, significantly more productive than hand labor – to complete the same production operation, it took machine labor, on average, just 17 percent of the time as hand labor (Atack, Margo, and Rhode 2022).

Although the productivity advantage of machine labor was substantial, regression analysis in Atack, Margo, and Rhode (2022) shows that only a surprisingly small percentage of the advantage – between a quarter and a third, depending on the estimation method -- can be attributed directly to the use of powered machinery. The main reason why is that about half of the production operations under machine labor did not use inanimate power, yet somehow workers performed these tasks more quickly than their counterparts did using hand labor methods. Further analysis showed that much of this higher productivity in the non-mechanized operations under machine labor can be attributed to other differences between the production methods, especially the greater use of division of labor. This greater use of division of labor in the factory is closely related to the concept of de-skilling. Our previous work, however, did not attempt to measure de-skilling directly. To show how this can be done, we first describe the collection and reporting of occupations in the HML study.

Occupations in the HML Study

Volume Two of the HML study reports the hand and machine production data at the operation level, providing a brief description of operations in the order in which they were performed; and a variety of additional information, including the occupational title of the worker and the time spent by the worker performing the operation. These data lie at the heart of our analysis in this paper. Specifically, we are interested in whether the time spent on the same operation in machine labor utilized less "skilled" labor than under hand labor. To operationalize this, we need to first classify workers by skill level.

We began by extracting a list of all unique occupation titles appearing in our digitized version of the HML study, regularizing their spelling, and making all singular. Where multiple occupations were reported, we adjusted accordingly. Thus, the occupation of "canner and labeler" was treated as two different occupations (with possibly different skill sets): "canner" and "labeler" whereas a "shank and toe nailer" was simply classified as a nailer (who happened to nail two different parts of a shoe together) and treated as one skill set.[9] Despite our best efforts, this remains almost certainly imperfect, as it was for census officials at the time.[10] For example, was a "cementer and channel layer" different from a "channel cementer"? We decided "yes" so

[9] When a person was listed with two or more occupations in HMLS, we ensured that there was a singular occupational entry for each occupation. If not, one was created. For example, "Shaft and Pole trimmer" was separated into "shaft trimmer" (of which there already were a number so identified) and "pole trimmer" (of which there were none prior to our parsing). The skill set used for classification is primarily defined by "trimmer".

[10] For example, "this indefiniteness in certain occupations makes their accurate return and classification impossible…[and] the great difficulty of securing absolutely correct returns from persons who are ignorant, indifferent, or not trained in making accurate statement, or who, as a matter of fact, do not know the precise nature of the occupations followed…" (United States and Hunt 1914, 5).

this affects our occupational counts but, since the skill levels were the same, it has no impact upon our deskilling metric.[11] We identified 1,703 different occupations in the HMLS after eliminating those occupations that were compounded by "and" but for which individual occupations were also recorded.[12]

We then used the work of Alba Edwards for the 1910 Census (United States. Bureau of the Census 1910; United States and Hunt 1914; United States and Hunt 1914; United States. Bureau of the Census., Hunt et al. 1915) and its expansions and elaboration (see Appendix A) to classify the various occupational titles into four categories—non-professional white collar; skilled blue collar (including foremen/forewomen); semi-skilled blue collar/operative; and common (unskilled) laborer. Because the HML study focused on the "shop floor" —that is, on the actual production of the good, rather than the management of the enterprise in which production took place – there were hardly any with occupations in the white-collar category although these were not totally absent (like proofreader, shipping clerk or bookkeeper). Where a particular person had two or more distinct occupations, we assigned them the higher skill rating. Thus, for example, the person whose occupation was given as "matcher and foreman" was assigned the supervisory code for a foreman rather than the operative code for a matcher. Our procedures are described in greater detail in Appendix A.[13]

---

[11] Similarly, we kept gendered occupational titles separate (for example, dairyman v. dairymaid).

[12] The "cementer and channel layer," for example, was instead recorded both as a "cemernter" and as a "channel layer" in our occupational statistics.

[13] For an earlier classification, see Hunt (1897). Hunt used four major categories: A. Proprietor Class --which included farmers; B. Clerical -- including foreman and weighers and measurers; C.

Once we had the occupational classifications and associated skill levels, we computed the occupational shares of production time under machine labor and hand labor. This is done using the HML crosswalk, and our associated concepts of an operation "block" and an operation "block link", as described in Atack, Margo, and Rhode (2022). Specifically, an operation block is a collection of production operations of size H (for hand labor) or M (for machine labor). H and M are non-negative integers and refer to the number of separate operations in the block such that the intermediate good entering and exiting the block is in the exact same stage of completion. A block link is a mapping, designated H:M, between the hand (H) and machine (M) blocks. We transcribed these mappings from the unit-specific tables in the report that were put together by the HML staff, that is, the HML crosswalk.

Table 1 shows the distribution of the block links in the regression sample that we use, along with relevant sample statistics. There are no 0:1 or 1:0 block links in the regression sample because the former represented hand operations that were no longer performed under machine labor, and the latter represented novel machine operations that were not performed under hand

---

Skilled Workers – including most manufacturing occupations; and D. Laborers. This mapping lumps most manufacturing workers, both skilled and semi-skilled, together. Edward's innovation was to break semi-skilled operative workers out from both skilled and unskilled raw labor in his so-called "social-economic" groups. In the "deskilling" debate, which focus on the shift away from craft-based production, such distinctions are crucial. Lumping craft-trained workers with operatives, as Hunt largely does, would hide key changes. Edward's treatment of operatives accords with Katz and Margo (2014, p. 17). It also better reflects the historical debate and our intent.

labor.[14] The block links that are relevant for our regression analysis are those that overlapped between the two methods—1:1, 1:M, H:1, and H:M—as these are the operations that represent the equivalent intermediate output, according by the HML staff. The table shows the average fractions using steam, water, or mechanized (that is, using steam and/or water) for each block link type under machine labor. As in Atack, Margo, and Rhode (2022), the mechanization measures are "one-touch" or extensive margin estimates—that is, if any operation within the machine block used inanimate power, we code the block link as "mechanized."

There are 4,405 block links in the sample that we use for our regression analysis (see the next section). Approximately 78 percent of these are 1:1 block links – that is singular tasks under hand labor that the HML staff also matched up to single tasks under machine labor. Approximately half (48.4 percent) of these blocks were mechanized under machine labor, using our one-touch measure.[15] The remaining 22 percent of the block links (N = 993) consisted of operations that underwent some degree of task reorganization and restructuring under machine labor compared with hand labor. Three-quarters (743/993) were 1:M or H:M, in which one or

[14] In the sample of units studied in this paper, there were 329 1:0 block links—hand operations no longer performed under machine labor—and 3,275 0:1 block links—novel machine operations. Many of these latter blocks, for example, were associated with steam power production (like "furnishing power"). Atack, Margo, and Rhode (2019) provides a detailed discussion of these block links framed within Acemoglu and Restrepo's (2018) model of automation.

[15] By "mechanized", we mean inanimately powered using a "one-touch" measure—that is to say, if any task within the block used inanimate power, the block itself was inanimately powered.

more hand operations were mapped into M machine operations. Overall, these were the most mechanized (79.9 percent) block links.[16] Although less common than 1:M or H:M block links, the H:1 links were also highly mechanized (74 percent). All told, about 55 percent of the block links were mechanized under machine labor.

Once we have the block links, we can compute the share of production time in the block performed by each category of labor, by production method. This allows us to compute the variable, *De-Skill*, as follows:

$$\textit{De-Skill} = 1\{\Delta \text{ (share of production time using semi-skilled operative + common labor)} > 0\}$$

where Δ is the difference between machine and hand labor. That is, *De-Skill* is a dummy variable taking the value 1 if the share of production time performed by semi-skilled operatives and common labor in the machine labor block exceeded the share in the corresponding block in the hand labor block; 0 otherwise.[17] We focus on a discrete measure of deskilling as this allows more direct, cleaner tests of the hypotheses.

[16] The figure is a weighted average of the mean mechanization rates of the 1:M and H:M block links.

[17] There is a small (less than 3 percent) of blocks for which we observe “up-skilling”. There are too few of these to analyze separately, and they are simply grouped with the 0 values of the dependent variable.

The final column of Table 1 shows the mean values of *De-Skill*, overall, and by block link type. The overall mean is 0.355 implying that, in approximately 36 percent of the block-links, relatively more use was made of common labor and operatives under machine labor than under hand labor – direct, unambiguous evidence of de-skilling in the traditional sense. It is also clear from the table that de-skilling was more frequent in the more complex block links compared with the 1:1 block links. As noted above, these more complex links were, on average, more mechanized, suggestive of a positive connection between greater use of powered machinery and de-skilling. The implications of this are addressed next through an econometric analysis of the block link data.

De-Skilling at the Operation Block Level: The Role of Mechanization

We estimate regressions of *De-Skill* using the sample of 4,405 block links that are matched between hand and machine labor through the HML crosswalk. Our goal is to measure the mean difference in *De-Skill* between the mechanized (inanimately powered) and non-mechanized operations within the same production units, controlling for various factors. We present OLS estimates first followed by an instrumental variable analysis. The discussion in the text focuses on our base specification (equation 1, immediately below).

OLS Estimation:

Our base regression specification is given by equation 1:

$$[Eq.\ 1]:\ De\text{-}Skill\ (a, j) = \beta(j) + \gamma(a) + \lambda * Mechanized\ (a,j) + \varepsilon\ (a,j)$$

The index *j* refers to the unit, and *a* to the block link. The *β(j)* are unit fixed effects. We include these because the manufactured goods produced by the various units were very different (such as circular saw blades vs. shoes) and there is no reason to believe that de-skilling would be the same across all products, controlling for the extent of mechanization. The *β(j)'s* differences such as the total number of workers in machine vs. hand labor, the year(s) to which the data pertain, and so on between the hand and machine labor establishments were the same for all block links within each unit. Because the unit fixed effects soak up all unit level differences, we cannot include any unit level differences in [eq. 1].[18] Our base specification also includes dummy variables for the block link types, γ(a), to control for the complexities introduced by multi-operation grouped tasks, which comprise about a quarter of the blocks.[19]

Table 2 reports the value of the coefficient λ for the single mechanization variable, *Mechanized*, which is the "one-touch" measure of mechanization of a machine labor block introduced in the previous section. While "mechanized" means that either steam or waterpower was used somewhere in the block, the overwhelming majority used steam. The reason for the single mechanization variable is discussed later in the paper. OLS identification of λ is achieved through the variation in *Mechanized* across block links within units. We expect to see that mechanization is positively associated with de-skilling, implying that $\lambda > 0$.

Why should λ be positive? The hypothesis is that mechanization promoted a greater degree of division of labor. As Joseph Roe (1926, 31) put it discussing the contributions of

---

[18] Later in the paper we substitute four-digit SIC industry codes for the unit fixed effects, which allows us to include unit level variables in the regression.

[19] The values of *β(j)* and γ(a) are available on request. The left-out block-link dummy is 1:1.

machines by Bentham and Brunel in the Portsmouth shipyard, "these machines were thoroughly modern in their conception and constituted a complete range of tools, each performing its part in a definite series of operations. By this machinery ten unskilled men did the work of 110 skilled workmen".

Economic theory provides support for Roe's argument. In Chaney and Ossa's (2013) model of a "production chain" – essentially analogous to the HML's ordering of production tasks -- the optimal division of labor is decreasing in the fixed costs to workers of acquiring the skills to perform various operations. All else equal, a reduction in fixed costs increases the profitability to the establishment of dividing up production tasks among more workers; in our view, use special-purpose inanimate powered machinery reduces the skills necessary to perform the tasks, thereby enabling greater division of labor (see also Becker and Murphy 1992).

Table 2 shows the OLS estimate of λ, 0.044, which is significant at the 5 percent level (s.e. = 0.018). The table also shows the percent of the mean value of *De-Skill* accounted for ("explained") by mechanization, which is the estimate of λ multiplied by the mean value of *Mechanized*, divided by the mean value of *De-Skill*. The percent explained is 6.8 percent [= ((0.044 x 0.552)/0.355) x 100 percent]. By construction, the remaining gap (93.2 percent) is explained collectively by the coefficients of the unit level and block link dummies.

Instrumental Variables Estimation

The results in Table 2 support the conventional wisdom that mechanization promoted de-skilling. However, while positive and statistically significant, the mechanization effect is quantitatively small, suggesting that mechanization *per se* may not have been a major factor behind de-skilling, a surprising result.

At this point it is worth asking if the OLS estimate of λ is affected by endogeneity bias. The HML staff did not randomly assign inanimate power and associated special purpose machinery to the specific machine labor operations that used them. That decision was made by someone in the original establishment, presumably an owner or manager, and presumably because this mechanization was expected to be profitable.

To assess the magnitude, if any, of endogeneity bias, we need an instrumental variable (IV) for *Mechanized* at the block link level. We follow the same identification strategy used in Atack, Margo, and Rhode (2022). This makes use of the textual descriptions of production operations appearing in the "General Table – Production by Hand and Machine Methods" in the column titled "Work Done," organized by unit number and production method, in Volume Two of the HML study, to construct an instrumental variable (see Figure 1).[20] We briefly describe the construction steps here (additional details can be found in Atack, Margo, and Rhode (2022)). First, we extracted all unique occurrences of gerunds appearing in the "Work Done" columns.[21] The first word in these descriptions is almost always a gerund describing the principal action

[20] As a follow-up exercise, in May 2023, we inquired of a new LLM model—ChatGTP—whether each of these activities "could be mechanized using the technologies of 1895." The results were not reliable as different inquiries regarding the same gerund produced different answer sand the model hallucinated into existence a series on automatic machines that did not, in fact, exist.

[21] A gerund is an English verb to which "-ing" has been appended. These function as a noun in grammatical context.

taking place in the operation, so we call this the "principal gerund."[22] Next, a member of our research term with expertise in the history of technology was given just the list of gerunds and asked to sort them into two bins without consulting the HML study.[23] Based solely on the expert's knowledge of the history, activities described by the gerunds where the expert believed there was some technical feasibility of mechanization worldwide by the end of the nineteenth century were sorted into one bin (bin #1), while those for which there was very little or none were sorted into the other (bin #0).

Table 3 (taken from Atack, Margo, and Rhode 2022) shows the distribution of the five most common principal gerunds for the hand blocks in the 1:1 and 1:M block links in the regression sample, grouped by bin #0 (little or no technical feasibility of mechanization) versus bin #1 (some technical feasibility). The five most common activities in bin #0 were "making", "putting", "overseeing", "finishing", and "marking." For each, human judgement played a substantial role, and the requirements of the activity were idiosyncratic. Conversely, the five

[22] Additional gerunds, if present, are always closely related to the main activity described by the principal gerund – the principal gerund is, in other words, the textual equivalent of a sufficient statistic.

[23] As discussed in Atack, Margo, and Rhode (2022), the list of gerunds was produced from a digitized version of the HML text. Because some gerunds can describe very different activities depending on a single letter – for example, "striping" vs. "stripping" – on rare occasions the expert was forced to consult the printed version of the HML study to be sure that the distinction was also present in the original text and not somehow garbled in the digitization, but the expert did not contemplate the text preceding or following the gerund in question.

most common activities in bin #1 were “cutting”, “sewing”, “smoothing”, “stitching”, and “conveying”. These are all repetitive activities for which special-purpose machinery had been invented in the nineteenth century.

As discussed in our earlier article, the motivation for the instrument derives from Acemoglu’s and Restrepo’s (2018) canonical model of automation. By the late nineteenth century, science and engineering had advanced to the point where the mechanization of certain physical activities – as captured by the relevant gerunds – was technically feasible. Conditional on the unit to be manufactured, therefore, the fraction of operations that, in principle, could be mechanized under machine labor depends on the classification of the relevant gerunds into the two bins. Variation in the classification corresponds to variation in the cost of mechanization; this variation is, by construction, exogenous, because it depends on technical feasibility, which is beyond the control of the owner or manager deciding on whether to mechanize or not, implying that our instrument satisfies the exclusion restriction.[24]

[24] For the 1:1 and 1:M block links there is a one-to-one mapping from the two bins to our IV, which is the “one-touch” analog to *Mechanization*, *MECHABLE* = 1 if principal gerund was sorted into bin #1, or 0 (if sorted into bin #0). For the H:1 and H:M block links there is an intermediate step in the construction of the IV because when H > 1 there may be more than one principal gerund. In the intermediate step we construct a weighted average of technical feasibility of mechanization for each principal gerund in the H operations in the hand block, where the weight is the share of time devoted to the operation in the overall time in the hand block. If the weighted average exceeds zero, MECHABLE = 1 for the overall block link. Note that, because we have one instrument, we can only have one endogenous variable (*Mechanized*).

Table 4 shows the results for the IV estimation of Eq. 1. The first stage coefficient of *MECHABLE*, 0.316 (s.e. = 0.020), is positive (as it should be) and the associated Kleibergen-Paap F-statistic (240.7) indicates that the instrument is very strong (p-value = 0.00001). The 2SLS estimate of λ, 0.096, is positive and significant (s.e. = 0.0.038) and about twice as large in magnitude as the OLS estimate.[25] If we repeat the decomposition exercise using the 2SLS estimate, the percent explained by mechanization is about 15 percent, compared with 7 percent for OLS.

The IV estimate indicates that the OLS bias is downward. Measurement error could cause a downward bias, but this is almost certainly not present here given the careful nature of the HML study and the fact that our measure of mechanization is "one-touch" – that is, we only require that the HML staff to have correctly observed the <u>use</u> of inanimate power. A more plausible explanation is reverse causality. In Atack, Margo, and Rhode (2022) we showed the OLS estimation of [eq. 1] produced upward bias in the estimate of λ when the dependent variable was the difference in production times between hand and machine labor. Our explanation is, when mechanization was a choice variable – as it clearly was in the establishments contributing their data to the HML study -- use of inanimate power was more likely when it was more profitable, and it was more profitable if the savings in production time was especially large. But if this is true, the share of production time that is mechanized under machine labor will be smaller than if mechanization were randomly assigned, an unobservable factor that will bias *De-*

[25] Note, however, that the 95 percent confidence interval around the 2SLS estimate of λ (0.021, 0.170), includes the OLS estimate so we cannot reject the hypothesis that the 2SLS and OLS estimates are the same.

Skill towards zero. Correcting for this bias results in a larger estimate of λ when IV is used instead of OLS.

De-Skilling: The Role of Factors Other than Mechanization

While mechanization did have a positive causal impact on de-skilling at the production operation level, most of the higher level of de-skilling under machine labor cannot be explained directly by greater mechanization. What, then, accounts for the unexplained portion?

To answer this question, we follow the strategy in Atack, Margo, and Rhode (2022), which uses the detailed descriptions of the goods produced to map the HML units into 70+ four-digit SIC codes. We substitute these SIC code dummies for the unit fixed effects, which results in equation 2.

$$[Eq.\ 2]: \Delta\ De\text{-}Skill = \eta(s) + \gamma(a) + \lambda^{*} Mechanized\ (a,j) + X(j)^{*}\gamma + \varepsilon\ (a,j)$$

The η(s) are coefficients of the 4-digit SIC fixed effects. Because eq. 2 does not include unit fixed effects, we can add continuous or dummy variables, X(j), at the unit level. Coefficients of the unit level variables are identified by variation across units within the 4-digit SIC codes. The unit variables of interest are not available for all units in the original regression sample, the sample size for eq. 2, about 3,900 block-links, is smaller than the sample size in Table 1 (see Atack, Margo et al. 2022).

OLS estimates of λ and β and the associated percent explained calculations are shown in Table 5. It is reassuring that the point estimate of λ, 0.049, is close that obtained using the eq. 1 specification with the Table 5 sample with unit fixed effects, suggesting that the combination of the 4-digit SIC dummies and particular unit level variables does an acceptable job of capturing the relevant variation in the data.[26]

Our main interest is the role of the division of labor. We include two measures of the division of labor – (1) the ln of the total number of production operations in manufacturing the unit and (2) the share of the number of operations performed by the average worker. The number of operations is available from the Volume One tables in the HML report. The Appendix B shows how we can combine unit level information in the report with the operations level data to construct the share measure. The hypothesis of interest is that greater overall division of labor led to more de-skilling, so we expect the coefficient of the first variable to be positive and of the second variable to be negative. We also control for two other "scale" variables, the number of different workers employed in producing the unit, and a "quantity production" dummy.27

Our main finding is that coefficient of the percent of tasks performed by the average worker is negative, highly significant, and large in absolute value, accounting for 40 percent of the mean value of *De-skill*. The mean value of this variable, -0.397, is negative, indicating that, under machine labor, the average number of operations performed by the average worker was

---

[26] See the notes to Table 5. We say "acceptable" because the adjusted-R square with 4-digit SIC codes (0.411) accounts for about 81 percent as much of the variance compared with the eq.1 specification for the same sample (adjusted R-square = 0.510).

[27] We cannot measure the division of labor at the block link level for all observations in the regression sample because to do so we would need the names of the individual workers which were not included in the published study. However, we can measure the division of labor at the unit level; see the Appendix B.

much smaller than under hand labor – that is, a far higher degree of division of labor. Once we control for the number of workers and the fraction of tasks performed by the average worker, the impact of an increase in the number of tasks to be performed is positive, but very small, and statistically insignificant.

We know that Carroll D. Wright, who guided the study in his role as the first U.S. Commissioner of Labor, considered a high degree of division of labor to be an essential feature of machine labor methods where "matters are so arranged that every workman has his particular work to perform, generally but a small portion of that which goes to the completion of the article to be produced" (U.S. Department of Labor. 1899, 1: 11). To the extent that some production tasks required less "skill" than others, a higher degree of division of labor should be associated with a higher level of de-skilling, but there is no inherent reason that the effect would be quantitatively large. The various nineteenth century American censuses of manufacturing never attempted to measure the division of labor directly, and its presence (and potential impact on productivity) in census data can only be inferred from variation across establishments in the number of workers. To our knowledge, the HML study is the only data source for nineteenth century US manufacturing for which direct measures of the division of labor can be constructed.

If we estimate the regression just with the number of workers, the coefficient is positive, highly significant, and large in magnitude, and by itself accounts for nearly half of the mean value of *De-Skill*. However, including the other three variables cuts the coefficient of the number of workers by nearly 60 percent in size, although it remains positive and statistically significant. The other scale variable, *Volume Production*, measures the difference between machine and hand labor using a dummy variable indicating whether the actual machine or hand quantity produced exceeded a critical cutoff level where substantial scale economies from the adoption of "flow" or "mass" production scheduling might have operated, leading to additional de-skilling (Hounshell 1984; Scranton 1997). [28]

Quantity production clearly was more common under machine labor, as indicated by the positive mean value (0.139) of the dummy variable. If, after controlling for the number of workers, quantity production *per se* enhanced de-skilling, the coefficient of the dummy variable should be positive. It is, 0.021, but the estimate is insignificant and its explanatory power in terms of accounting for the mean value of *De-Skill* is negligible. This suggests that the number of workers, once we control for the fraction of tasks performed by the average worker, is mainly capturing additional effects of a greater division of labor on *De-Skill*, rather than larger volume production.

The HML staff also collected data on average daily hours. Over the nineteenth century, there was a downward trend in average daily hours; on average, daily hours were shorter in the machine labor establishments (Atack and Bateman 1992; Atack, Bateman et al. 2003). If less skilled workers were able to maintain work effort continuously for longer periods of time, we would see a positive association between average daily hours and de-skilling – and the coefficient, indeed, is positive. However, the explanatory power of the variable is slight and, in fact, negative, because average daily hours was shorter under machine labor.

[28] While the ideal cutoff should be guided by historical examples and discussion, the relevant literature provides no operational guidance. The threshold we adopted was the quantity associated with the 75th percentile of the machine labor distribution, 1,500; see Atack, Margo, and Rhode (2022).

The HML staff attempted to assess the quality differences in goods produced by the machine and hand units in the surveyed establishments. For the most part, they concluded there was either no meaningful difference or a difference in favor of the machine labor version of the good. That said, if the hand product were of better quality, we might expect a greater level of skill involved in its construction, and hence a higher value of de-skilling when the good was made by machine labor. The coefficient is positive, but insignificant, and its explanatory power in accounting for the mean value of *De-Skill* is very small.

We include the difference in the observation year between machine and hand labor in the regression. This difference is almost always positive, and the mean is 27, indicating that, on average, the hand labor data were older than the machine labor data by 27 years. The coefficient of this variable is negative, 0.015, and significant at the 10 percent level (s.e. = 0.010), indicating that the older the hand labor data was relative to machine labor, the greater the relative use of skilled labor in hand production of the unit, other factors held constant. This pattern makes sense because earlier hand labor establishments were likely closer to the "old-fashioned" methods that Wright had in mind for the study.

We cannot and do not claim that the coefficients of these unit level variables reflect causal impacts, and we have no way of instrumenting them individually. Still, the mere fact that we can include these measures at all in the regression goes far beyond what is possible with other nineteenth century data. All told, the unit level variables account for 74 percent of the mean value of *De-Skill*, compared with 6 percent for mechanization per se.

Concluding Remarks

Economic historians have long believed that de-skilling occurred during the first industrial revolution but have been largely stymied in measuring its extent or determining if the diffusion of inanimate power was primarily responsible. In the American case, the key problem is that the US manufacturing censuses during the nineteenth century did not measure de-skilling directly; while it is possible to document the occupations of manufactured workers using census data, it is not possible to link the occupations to data on use of inanimate power.

This paper has turned to an entirely different source, the Department of Labor's 1899 *Hand and Machine Labor* study. The HML study allows us to measure de-skilling directly at the production operation level and relate it to the use of inanimate power. We find de-skilling was quite common in the mechanized factory of the late nineteenth century compared with traditional artisan hand production, and that greater use of inanimate power led to greater de-skilling. However, the magnitude of the mechanization effect was relatively modest.

De-skilling was part and parcel of a vast growth in the division of labor over the nineteenth century. While we cannot measure the change in division of labor continuously over the nineteenth century, it is clear from the HML study that this change was strongly correlated with growth in establishment size, where size is measured by the number of workers. Our regression analysis of the HML data show that de-skilling was strongly increasing in the number of workers, in main part because the division of labor was increasing in the number of workers.

The HML study did not investigate why the division of labor increased in the transition from hand to machine labor but there is little doubt that the transportation revolution was a critical factor (Atack, Haines et al. 2011). The transportation revolution increased market access and in so doing, made a larger scale of operation more profitable – as the saying (from Adam

Smith) goes, “the division of labor is limited by the extent of market”. As the division of labor increased, workers became more specialized in production, and the “average worker” was a convex combination of individuals performing different operations according to comparative advantage, more productive than a single artisan performing all tasks from start to finish. Compared with such artisans, the typical nineteenth century factory operative had much less to learn on the job, lowering the costs of supplying labor to manufacturing. Although advances in technology and emerging complementarity with capital increased the skill demands on factory workers, this calculus remained the same until well into the twentieth century when the forces of automation eventually caught up, making operatives highly vulnerable to displacement by machinery (Goldin and Katz 1998; Acemoglu and Restrepo 2018).

Table 1: Distribution and Sample Statistics by Block-Link Type: Regression Sample

Block Link Type (**H**and:**M**achine)	Number of Block Links	Mean Fraction Steam, Machine Labor	Mean Fraction Water, Machine Labor	Mean Fraction Mechanized, Machine Labor	Mean Value, *De-Skill*
1:1	3,412	0.460	0.025	0.484	0.323
1:M, M > 1	619	0.732	0.055	0.784	0.489
H:1, H > 1	250	0.704	0.052	0.744	0.416
H:M, H, M> 1	124	0.815	0.073	0.879	0.460
Total, Regression Sample	4,405	0.522	0.032	0.552	0.356

Source: computed from digitized HML study (U.S. Department of Labor. 1899).

Notes: Block links are defined as follow -- 1:1: a single hand labor operation is mapped to a single machine labor operation; 1:M, M> 1: a single hand labor operation is mapped to a block of M machine operations, M> 1; H:1, H> 1: A block of H (>1) hand operations is mapped to a single machine labor operation; H:M: A block of H hand labor operations is mapped to a block of M machine labor operations, H and M > 1. Mechanized = 1 if machine block used steam or waterpower or both; see text. NA: not applicable.

Table 2: OLS Regression and Decomposition Analysis of *De-Skill*

Variable	Coefficient	Mean Value of Independent Variable	Percent Explained of Mean Value of *De-Skill*
Mechanized	0.044 (0.018)	0.552	6.7%
Adjusted R-Square	0.670		

Source: see Table 1.

Notes: The decomposition ("Percent Explained of Mean Value of *De-Skill*") is computed by multiplying the regression coefficient of *Mechanized* (0.044) by the mean value of *Mechanized* (0.552) and dividing the product (0.024) by the mean value of *De-skill* (0.356) = (0.024)/(0.356) = 0.067, or 6.7 percent. The sample size for the regression is 4.,405 block links. The regression also includes dummy variables for block link types and for units (see the text). Standard errors are clustered at the unit level. Sample means are from Table 1.

Table 3: The Top-Five Activities (Principal Gerunds) in Hand Production, 1:1 and 1:M Block-Links in the Regression Sample: Bin #0 versus Bin #1

Bin #0	Number	Bin #1	Number
Making	156	Cutting	597
Overseeing	141	Sewing	136
Putting	121	Smoothing	77
Finishing	38	Stitching	76
Marking	34	Conveying	73
Total count in Bin #0	690	Total count in Bin #1	2,968

Source: see Table 1 and Atack, Margo, and Rhode (2022).

Notes: The table shows the distribution of the five most common principal gerunds classified into bin #1 (some feasibility of mechanization) versus bin #0 (little or no feasibility of mechanization) for the 1:1 and 1:M block links in the regression sample.

Table 4: 2SLS Regression and Decomposition Analysis of Δ *De-Skill*

Variable	Coefficient	Mean Value of Independent Variable	Percent Explained of Mean Value of *De-Skill*
First stage, *MECHABLE*	0.315 (0.020)	0.820	
Kleibergen-Paap F-Statistic	240.7		
2SLS, *Mechanized*	0.096 (0.038)	0.552	14.9%

Notes: The decomposition ("Percent Explained of Mean Value of *De-Skill*") is computed by multiplying the 2SLS regression coefficient of *Mechanized* (0.096) by the mean value of *Mechanized* (0.552) and dividing the product (0.053) by the mean value of De-Skill (0.356) = (0.053)/(0.356) = 0.149, or 14.9 percent. The sample size is 4,405 block links. The first stage and 2SLS regressions also includes dummy variables for block link types and for units (see the text). The Kleibergen-Paap F-statistic refers to the instrumental variable *MECHABLE.* Standard errors are clustered at the unit level. Sample means of *De-Skill* and *Mechanized* are from Table 1.

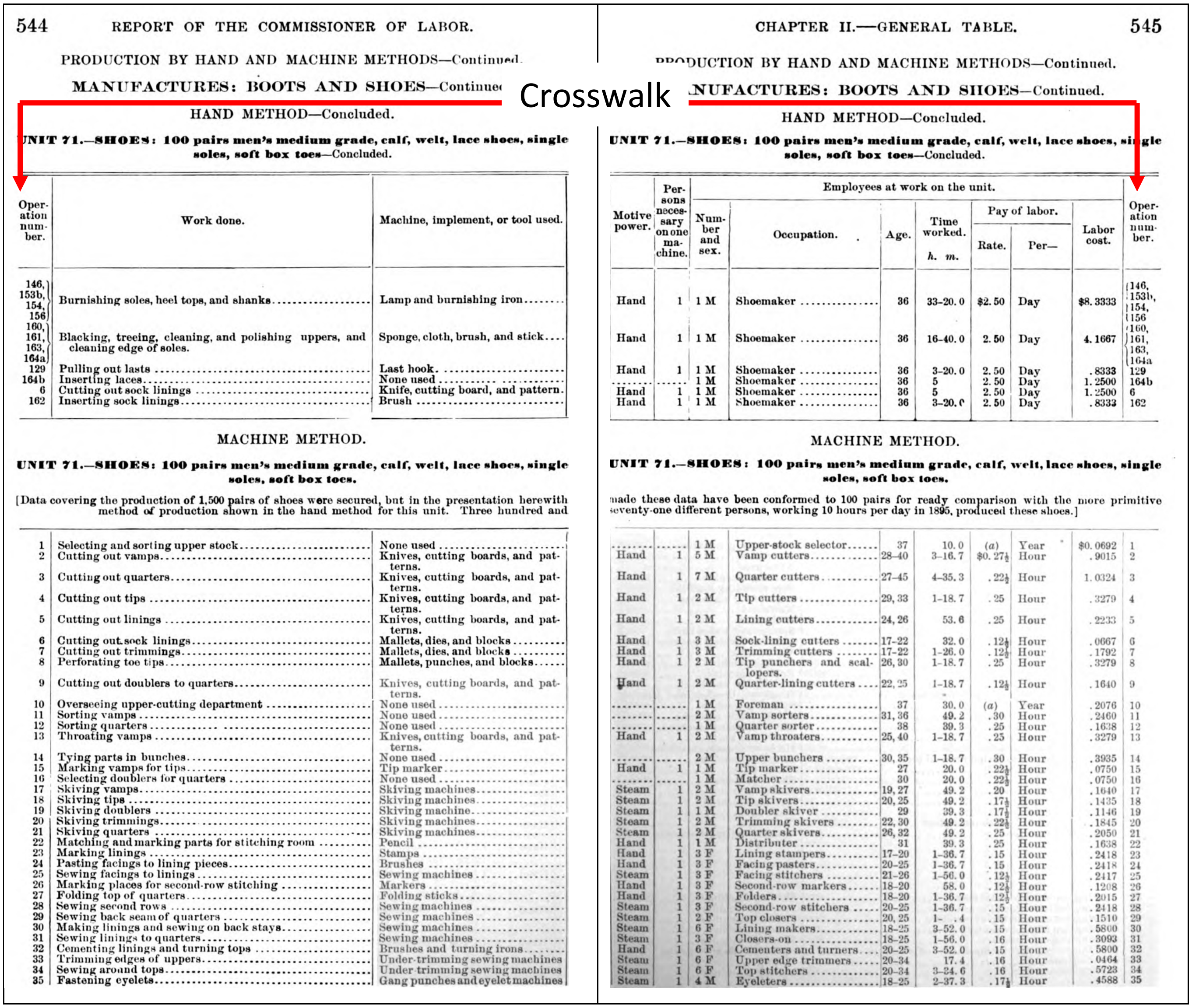
Table 5: OLS Regression of Δ *De-skill* with 4-Digit SIC Code Dummies

Variable	Coefficient	Coefficient	Mean Value of Independent Variable	Percent Explained at Sample Mean of Δ *Ln T*
Mechanized	0.052 (0.019)	0.040 (0.018)	0.554	5.8%
Unit Level Difference, Machine – Hand Labor:				
Ln (# of Workers)	0.103 (0.016)	0.041 (0.014)	1.848	19.9
Ln (# of Operations)		0.018 (0.022)	0.472	2.2
Fraction of Total Operations Performed by the Average Worker		-0.399 (0.069)	-0.397	41.6
Volume Production	-0.016 (0.027)	0.021 (0.029)	0.138	0.8
Ln (Daily Hours of Operation)	-0.0004 (0.189)	0.166 (0.175)	-0.030	-1.3
Hand Unit Quality Better	-0.005 (0.043)	0.054 (0.054)	0.030	0.4
Year of Observation	0.003 (0.010)	0.015 (0.010)	2.7	10.6
Percent Explained by Unit Level Variables				74.2
Percent Explained, Unit Level + Percent Mechanized				80.0
Adjusted R-Square	0.539	0.566		

Source: see Table 1.

Notes: Sample size is 3,876 block links because of missing data on some unit level variables (see text). Mean value of Δ *De-Skill* for this sample is 0.554. Regression includes 4-digit SIC code dummies and block link dummies (coefficients not reported). Coefficient of *Mechanized* = 0.049 (s.e. = 0.020) with SIC code and block-link dummies. Volume Production is the difference between two dummy variables, (Volume_Machine – Volume_Hand). Volume_Machine (Volume_Hand) = 1 if actual quantity>1,500; 0 otherwise. Hand Better = 1 if HML staff judged the product made by the hand labor establishment to be of better quality than the machine labor establishment. Standard errors are clustered at the unit level.

Figure 1: Verso and Recto Pages of part of the Hand and Machine Labor Table for Unit 71
Showing Raw Data and the HML-generated Crosswalk



544 REPORT OF THE COMMISSIONER OF LABOR.

CHAPTER II.—GENERAL TABLE. 545

PRODUCTION BY HAND AND MACHINE METHODS—Continued.

MANUFACTURES: BOOTS AND SHOES—Continued.

Crosswalk

HAND METHOD—Concluded.

UNIT 71.—SHOES: 100 pairs men's medium grade, calf, welt, lace shoes, single soles, soft box toes—Concluded.

Operation number.	Work done.	Machine, implement, or tool used.	Motive power.	Persons necessary on one machine.	Number and sex.	Occupation.	Age.	Time worked. h. m.	Pay of labor. Rate.	Pay of labor. Per—	Labor cost.	Operation number.
146, 153b, 154, 156	Burnishing soles, heel tops, and shanks	Lamp and burnishing iron	Hand	1	1 M	Shoemaker	36	33-20.0	$2.50	Day	$8.3333	146, 153b, 154, 156
160, 161, 163, 164a	Blacking, treeing, cleaning, and polishing uppers, and cleaning edge of soles.	Sponge, cloth, brush, and stick	Hand	1	1 M	Shoemaker	36	16-40.0	2.50	Day	4.1667	160, 161, 163, 164a
129	Pulling out lasts	Last hook	Hand	1	1 M	Shoemaker	36	3-20.0	2.50	Day	.8333	129
164b	Inserting laces	None used	....	......	1 M	Shoemaker	36	5	2.50	Day	1.2500	164b
6	Cutting out sock linings	Knife, cutting board, and pattern.	Hand	1	1 M	Shoemaker	36	5	2.50	Day	1.2500	6
162	Inserting sock linings	Brush	Hand	1	1 M	Shoemaker	36	3-20.0	2.50	Day	.8333	162

MACHINE METHOD.

UNIT 71.—SHOES: 100 pairs men's medium grade, calf, welt, lace shoes, single soles, soft box toes.

[Data covering the production of 1,500 pairs of shoes were secured, but in the presentation herewith made these data have been conformed to 100 pairs for ready comparison with the more primitive method of production shown in the hand method for this unit. Three hundred and seventy-one different persons, working 10 hours per day in 1895, produced these shoes.]

Operation number.	Work done.	Machine, implement, or tool used.	Motive power.	Persons necessary on one machine.	Number and sex.	Occupation.	Age.	Time worked. h. m.	Pay of labor. Rate.	Pay of labor. Per—	Labor cost.	Operation number.
1	Selecting and sorting upper stock	None used	....	......	1 M	Upper-stock selector	37	10.0	(a)	Year	$0.0692	1
2	Cutting out vamps	Knives, cutting boards, and patterns.	Hand	1	5 M	Vamp cutters	28-40	3-16.7	$0.27½	Hour	.9015	2
3	Cutting out quarters	Knives, cutting boards, and patterns.	Hand	1	7 M	Quarter cutters	27-45	4-35.3	.22½	Hour	1.0324	3
4	Cutting out tips	Knives, cutting boards, and patterns.	Hand	1	2 M	Tip cutters	29, 33	1-18.7	.25	Hour	.3279	4
5	Cutting out linings	Knives, cutting boards, and patterns.	Hand	1	2 M	Lining cutters	24, 26	53.6	.25	Hour	.2233	5
6	Cutting out sock linings	Mallets, dies, and blocks	Hand	1	3 M	Sock-lining cutters	17-22	32.0	.12½	Hour	.0667	6
7	Cutting out trimmings	Mallets, dies, and blocks	Hand	1	3 M	Trimming cutters	17-22	1-26.0	.12½	Hour	.1792	7
8	Perforating toe tips	Mallets, punches, and blocks	Hand	1	2 M	Tip punchers and scallopers.	26, 30	1-18.7	.25	Hour	.3279	8
9	Cutting out doublers to quarters	Knives, cutting boards, and patterns.	Hand	1	2 M	Quarter-lining cutters	22, 25	1-18.7	.12½	Hour	.1640	9
10	Overseeing upper-cutting department	None used	....	......	1 M	Foreman	37	30.0	(a)	Year	.2076	10
11	Sorting vamps	None used	....	......	2 M	Vamp sorters	31, 36	49.2	.30	Hour	.2460	11
12	Sorting quarters	None used	....	......	1 M	Quarter sorter	38	39.3	.25	Hour	.1638	12
13	Throating vamps	Knives, cutting boards, and patterns.	Hand	1	2 M	Vamp throaters	25, 40	1-18.7	.25	Hour	.3279	13
14	Tying parts in bunches	None used	....	......	2 M	Upper bunchers	30, 35	1-18.7	.30	Hour	.3935	14
15	Marking vamps for tips	Tip marker	Hand	1	1 M	Tip marker	27	20.0	.22½	Hour	.0750	15
16	Selecting doublers for quarters	None used	....	......	1 M	Matcher	30	20.0	.22½	Hour	.0750	16
17	Skiving vamps	Skiving machines	Steam	1	2 M	Vamp skivers	19, 27	49.2	.20	Hour	.1640	17
18	Skiving tips	Skiving machines	Steam	1	2 M	Tip skivers	20, 25	49.2	.17½	Hour	.1435	18
19	Skiving doublers	Skiving machine	Steam	1	1 M	Doubler skiver	29	39.3	.17½	Hour	.1146	19
20	Skiving trimmings	Skiving machines	Steam	1	2 M	Trimming skivers	22, 30	49.2	.22½	Hour	.1845	20
21	Skiving quarters	Skiving machines	Steam	1	2 M	Quarter skivers	26, 32	49.2	.25	Hour	.2050	21
22	Matching and marking parts for stitching room	Pencil	Hand	1	1 M	Distributer	31	39.3	.25	Hour	.1638	22
23	Marking linings	Stamps	Hand	1	3 F	Lining stampers	17-20	1-36.7	.15	Hour	.2418	23
24	Pasting facings to lining pieces	Brushes	Hand	1	3 F	Facing pasters	20-25	1-36.7	.15	Hour	.2418	24
25	Sewing facings to linings	Sewing machines	Steam	1	3 F	Facing stitchers	21-26	1-56.0	.12½	Hour	.2417	25
26	Marking places for second-row stitching	Markers	Hand	1	3 F	Second-row markers	18-20	58.0	.12½	Hour	.1208	26
27	Folding top of quarters	Folding sticks	Hand	1	3 F	Folders	18-20	1-36.7	.12½	Hour	.2015	27
28	Sewing second rows	Sewing machines	Steam	1	3 F	Second-row stitchers	20-25	1-36.7	.15	Hour	.2418	28
29	Sewing back seam of quarters	Sewing machines	Steam	1	2 F	Top closers	20, 25	1- .4	.15	Hour	.1510	29
30	Making linings and sewing on back stays	Sewing machines	Steam	1	6 F	Lining makers	18-25	3-52.0	.15	Hour	.5800	30
31	Sewing linings to quarters	Sewing machines	Steam	1	3 F	Closers-on	18-25	1-56.0	.16	Hour	.3093	31
32	Cementing linings and turning tops	Brushes and turning irons	Hand	1	6 F	Cementers and turners	20-25	3-52.0	.15	Hour	.5800	32
33	Trimming edges of uppers	Under-trimming sewing machines	Steam	1	6 F	Upper edge trimmers	20-34	17.4	.16	Hour	.0464	33
34	Sewing around tops	Under-trimming sewing machines	Steam	1	6 F	Top stitchers	20-34	3-24.6	.16	Hour	.5723	34
35	Fastening eyelets	Gang punches and eyelet machines	Steam	1	4 M	Eyeleters	18-25	2-37.3	.17½	Hour	.4588	35

Source: (U.S. Department of Labor. 1899, 2: 820-1)

Appendix A: Coding Occupations

In preparation for the Thirteenth Census in 1910, the Bureau of the Census had undertaken an extensive review of the occupational returns from 1900 for several states eventually extracting some 16-18,000 different occupational titles. These were then culled to about 13,000 although the Bureau opined "it is safe to estimate that, on the whole, there are at least two designations for each occupation, and hence the index does not contain over 7,000 or 8,000 separate and distinct occupations." This culled list would be provided to enumerators for the 1910 Census to serve as a guide although it was deemed "far from being all-inclusive, and will need to be supplemented by the addition of the new occupations returned in 1910" (United States. Bureau of the Census 1910, iii).

Subsequently, the Bureau produced a large Census volume on the Occupation Statistics (United States and Hunt 1914) along with a shorter summary Bulletin on the same (United States and Hunt 1914), observing "the value of occupation statistics is dependent very largely upon the form in which the occupations are classified" (United States and Hunt 1914, 17). The Bureau would follow up on this observation beginning with a monograph "Index to Occupations: Alphabetical and Classified" written by Alba Edwards who categorized the occupations into 215 main occupations and occupation groups, 84 of which were further subdivided to produce 428 separate occupations and occupation groups (U.S. Department of Commerce. Bureau of the Census, Hunt et al. 1915). Edwards (1917, 645 and Table 1) subsequently regrouped and refined these into 9 groups which he termed "Social-Economic Groups":

1. proprietors, officials, and managers;
2. clerks and kindred workers;
3. skilled workers;

4. semiskilled workers;

5. laborers;

6. servants;

7. public officials;

8. semiofficial public employees;

9. professional persons.

Since the HML study focuses solely upon manufacturing production, all workers in it fall into Edwards' categories 1-5.

We have chosen to further collapse these groups down to four by treating those businesses with just a solitary worker—their proprietor—as an employee/skilled worker rather than proprietor. Thus, for example, the shoemaker—the only worker in the Hand method of Unit 71 (see Figure 1) was assigned to group 3 of Edward's classification rather than to group 1. Almost all workers identified in the HML study fall into groups 3-5 although a small number with occupations like bookkeeper, counter and clerk were classified into group 2.

Two of Edwards' occupational categories explicitly contain the term "skill" although Edwards definition of the word is very narrow and specific: "properly applied only to those occupations in which the expenditure of muscular force is one of the chief characteristics. Within this field, those occupations have been considered skilled … [requiring] … a long period of training, or an apprenticeship … and … a degree of judgment, and manual dexterity". Extending this line of argument, he classified "semiskilled" occupation as those for which "only a short period or no period of preliminary training is necessary, and which in their pursuance call for only a moderate degree of judgment or of manual dexterity." He classified as "Laborers" or "unskilled" occupations requiring "no special training, judgment, or manual dexterity, but supply mainly muscular strength in the performance of course, heavy work," that is, by inference, the lacked skill (Edwards 1917, 646).[1] Clerical workers, professionals, and the like had other skills and attributes unrelated to the physical demands of their work. They would, in time, be referred to as "white collar" workers to indicate that their clothes would remain unsoiled by their labor.[2]

Over the years, Edwards' skill-based groups evolved and were further refined. For example, the Department of Labor's Employment Service prepared a dictionary of occupational titles with the objective of establishing broad groups of occupations requiring similar skills and abilities (U.S. Department of Labor. Employment Service 1939). It was followed by an alphabetical index of occupations and industries for the Sixteenth Census (1940) again authored by Edwards (United States. Department of Commerce. Bureau of the Census. and Edwards 1940) and similar volumes were produced to accompany subsequent censuses.[3]

In particular, the index prepared for the 1950 Census {United States. Department of Commerce. Bureau of the Census., 1950 #2277, also available on-line at https://usa.ipums.org/usa/resources/volii/Occupations1950.pdf} was used by IPUMS to generate their "OCC1950" variable found in all IPUMS census samples from 1850 onwards (https://usa.ipums.org/usa-action/variables/OCC1950#codes_section). Moreover, the characteristics identified by Edwards (training, dexterity, exercise of judgment, physical

---

[1] This categorization also apparently stood the test of time, see Edwards (1938, p. 3)

[2] Wikipedia credits the term "white collar" to Upton Sinclair, beginning in the 1930s. The phrase does not appear in Edwards' work so far as we can tell.

[3] See, for example, https://www.census.gov/topics/employment/industry-occupation/guidance/indexes.html (visited May7, 2023).

strength) were subsequently elaborated and extended by the U.S. Bureau of Employment Security in their detailed survey of the skills, training, and personal characteristics based on a sampling of workers performing 4,000 jobs (United States. Department of Labor. Bureau of Employment Security. 1956).[4]

We sought to link the HMLS occupational titles in a similar manner to standard skill classifications. This proved challenging because many of the occupation titles in the HMLS were highly specific to the production processes described (and indeed many were closely related to the gerunds defining the task activities). Another problem was that some titles were common across units and industries but referred to fundamentally different activities.

The first problem derived from extensive sub-division of activities to the task level by the HMLS; the second problem from lumping (or at least linguistic lumping) of similar actions and activities but involving, say, different materials (e.g. metal v wood).

Our initial attempts to mechanically link either to the published census or to the occupational text from the manuscript micro sampled data proved infeasible. In the former case, the HMLS has far more occupational titles than the published census and displays a wider range of skill requirements. In the latter case, the textual descriptions of occupations as recorded by the original enumerators and as transcribed by later data entry personnel was often incomplete even when almost certainly referring to the same specific occupation.

Consequently, instead of using mechanical linkage methods, the research team used their expertise to locate the titles in occupational lists/dictionaries, particularly the U.S. Department of

[4] For a use of these latter data in the context of tasks, focusing on the early twentieth century when electricity was replacing the steam engine as the prime mover, see Gray (2013).

Labor's *Dictionary of Occupational Titles* which sought to define the activities associated with specific jobs, much of it based upon direct observation or job analyses (U.S. Department of Labor. Employment Service 1939, iii).[5] They then assigned what they determined was the best match code using the 1950 codes (United States. Department of Commerce. Bureau of the Census. 1950) to the HMLS occupational titles. In assigning these codes, like the IPUMS coding with "OCC1950," we endeavored to match skill level and function over time. Thus, for example, a stagecoach driver or a wagon driver was classified as "bus and truck drivers" in "OCC1950".

To determine the appropriate 1950 occupational code, we initially created a Pivot Table in Excel on the transcribed occupations in the HML tables. This provided us with a list and count of the unique occupational titles and enabled us to identify plurals (and remove), spelling variations, and to separate and reallocate those workers with two or more jobs lumped together by "and" into the singular versions of each. This generated a list of 1,703 different occupational titles in the HMLS.

Two team members completed the coding task independently. Based on broad metrics, their initial assignments agreed in about 94 percent of cases. The residual 6 percent of disagreements were resolved quickly, without issue.

[5] This was the first such dictionary for U.S, occupations. The British had developed one similar for their industries and country-specific terminology more than a decade earlier (Great Britain. Ministry of Labor 1927).

The most common occupation was “machine hand” (coded as semi-skilled: 6900). The ten most common unique occupations that we identified along with their respective share of the over 22,000 tasks reported are shown in Table A.1.[6]

Where the term “assistant” or “helper” was appended to the search term, we treated these individuals as someone with the same skill set as specified in the occupational title but who had not yet been assigned full responsibility and independence as a skilled worker, placing them in IPUMS 4-digit coding for Apprentices, 6020-6150. These are persons in-training to be skilled craftsmen and thus treated as semi-skilled. This scheme differs somewhat from that used in the 1950 Alphabetical Index (United States. Department of Commerce. Bureau of the Census. 1950). where helpers are classified as 690 (=operatives rather than apprentices). Both, however, are coded as semi-skilled which is the important result.

As a final cross-check, we also compared our occupational classifications against the wage distribution to identify potential problematic cases requiring further review, but we did not directly use wages to assign skill classifications.

[6] The apparent importance of “watchmaker” reflects the subdivision of watchmaking tasks in the HMLS data rather than the actual number of watchmakers.

Table A.1
The Ten Most Common Occupations by Task in the Hand and Machine Labor Study

Occupation	Share of Tasks reporting this Occupation
Machine hand	5.4%
Foreman	4.5%
Watchmaker	3.1%
Engineer	2.5%
Fireman	2.4%
Cutter	2.3%
Blacksmith	1.9%
Sawyer	1.9%
Shoemaker	1.8%
Laborer	1.5%

Appendix B: Measuring the Division of Labor in the HML Study

Compared with other sources on nineteenth century manufacturing, such as the Census of Manufactures, a unique feature of the HML study is that it provides direct evidence on the division of labor. This is because the study reported information about each production operation in the manufacturing of the unit, including the number of workers performing the operation. In addition, the study reported the number of different workers employed in producing the unit. With these two pieces of information, we can construct two unit-level measures of the division of labor – the number of distinct operations involved in producing the unit and the proportion of operations performed by the average worker.

To fix ideas, let X be the quantity of the good to be produced, S = {s_1, …, s_n} be the list of n operations involved in producing X, and S* to consist of all proper subsets of S except for the empty set, plus S itself. Initially, we take S as given – that is, we abstract from any economic decision involved in dividing overall production into more or fewer than n operations.

Next, we define D to be a mapping from S* to W, D(S*) → W, where is a list of the workers employed to produce X. We say that "division of labor" has occurred if (i) every operation in S appears in at least one element in the mapping (ii) at least one element in the mapping is a proper subset of S. The first requirement is the essentiality requirement – to produce X, each operation in S must be performed by some worker. The second requirement rules out that every worker is mapped to S itself, which would mean that no worker specializes in a subset of S.

For example, suppose N = 4 (that is, four operations) must be performed to produce X, and suppose there are two workers. In this case, there are 15 subsets in S*, one of which is S. Suppose that D = {(1, 2), (3, 4)}, so that worker #1 performs the first two operations (1, 2) and the second worker performs the final two (3, 4). D is a valid mapping because all operations are performed and neither worker is mapped to S itself, so division of labor has occurred. Conversely, suppose that worker #1 performs (1, 2, 3, 4) and so, too, does worker #2. In this case, both workers replicate each other; no division of labor has occurred. An intermediate example would be worker #1 performing (1, 2, 3) and worker #2 performing (3, 4) – both workers perform the third operation, but this is the only one in common between them.

As the examples in the previous paragraph illustrate, in any valid mapping D – that is, when division of labor is occurring – the extent of <u>specialization</u> across workers may be <u>incomplete</u> or <u>complete</u>. It is incomplete if there is some overlap in the tasks performed by different workers. In the final example, imagine that the penultimate task is "finishing" and the final task is "packaging". The first worker performs operations #1 through #3, while the second performs operations #3 and #4, so both do finishing (operation #3) but only the second worker does packaging (#4). By contrast, specialization is complete if there is no overlap of tasks – as in the first example, where the first worker performs operations #1 and #2 and the second worker performs #3 and #4.

For any given S, there is a maximal degree of specialization in which each worker performs one and only one task. A necessary (but not sufficient) condition for maximal specialization is that the number of distinct workers be at least as large as N, the number of tasks in S. However, there may be tasks which require more than one worker – for example, because a machine is used that requires two operators – or there may be tasks which are replicated to generate the requisite flow of production. To continue with the previous example, suppose that there are now 6 workers. Worker #1 performs operation #1 and worker #2 performs operation #2; workers #3 and #4 perform operation #3 and workers #5 and #6 perform operation #4. Even though the number of workers exceeds the number of operations, specialization is maximal because the mapping D is still one to one, except for the intensive margin on operations #3 and #4. This intensive margin might reflect the nature of the machines used in tasks #3 and #4, which may require more than one worker to operate properly.

With the above discussion in mind, how might the division of labor be quantified? As a step towards answering this question note that we can represent M by a matrix whose elements m(j, k) index workers (j) and operations (k) to which they are assigned. A variety of functions defined over the m (j, k) could serve as summary statistics of the division of labor. The statistics could be equally weighted across workers or weighted by the amount of time needed to complete the operation or the degree of skill required to perform it.

To construct such statistics, however, we must be able to construct the mapping M – that is, allocate specific workers to specific tasks. As previously noted, although the HML survey form did identify which workers performed which operations, this level of detail was suppressed in the published study, and the original survey forms are no longer extent. Despite this limitation, it is possible to compute an <u>average</u> measure of the division of labor for the overall unit. This statistic is P(n), the proportion of operations performed by the average worker:

$$P(n) = ([\sum \text{(number of workers assigned to operation k)}] / \text{Total number of different workers})/n$$

To see this statistic in action, consider the above example where n = 4, there are two distinct workers, the first performing tasks #1-3 and the second performing tasks #3-4, and both performing tasks #9 and #10. In this example, the summand in the numerator will be 5; this is divided by 2, the number of distinct workers, and then further divided by 4, the number operations. The result is 0.625 – the average worker performs 63 percent of the operations in producing the good. In the econometric analysis in the final section of the paper, we use P(n) and n as the two measures of the division of labor.

Bibliography

Acemoglu, D. and P. Restrepo (2018). "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." American Economic Review **108**(6): 1488-1542.

Atack, J. and F. Bateman (1992). "How Long Was the Workday in 1880?" Journal of Economic History **52**(1): 129-160.

Atack, J., F. Bateman and R. A. Margo (2003). "Productivity in manufacturing and the length of the working day: evidence from the 1880 census of manufactures." Explorations in Economic History **40**(2): 170-194.

Atack, J., F. Bateman and R. A. Margo (2008). "Steam Power, Establishment Size, and Labor Productivity in Nineteenth Century American Manufacturing." Explorations in Economic History **45**(April): 185-198.

Atack, J., F. Bateman and T. Weiss (1980). "The Regional Diffusion and Adoption of the Steam Engine in American Manufacturing." Journal of Economic History **40**(2): 281-308.

Atack, J., M. Haines and R. A. Margo (2011). Railroads and the Rise of the Factory: Evidence for the United States, 1850-1870. Economic Evolution and Revolutions in Historical Time. P. W. Rhode, J. Rosenbloom and D. Weiman. Palo Alto, Stanford University Press**:** 162-179.

Atack, J., R. A. Margo and P. W. Rhode (2019). ""Automation" of Manufacturing in the Late Nineteenth Century: The Hand and Machine Labor Study." Journal of Economic Perspectives **33**(2): 51-70.

Atack, J., R. A. Margo and P. W. Rhode (2022). "'Mechanization Takes Command?': Powered Machinery and Production Times in Late Nineteeth-Century American Manufacturing." Journal of Economic History **82**(3): 663-690.

Atack, J., R. A. Margo and P. W. Rhode (2023). "Digitizing Carroll D. Wright's "Hand and Machine Labor" Study." Jahrbuck Fur Wirtschaftsgeschichte **64**(1): 9-30.

Becker, G. S. and K. M. Murphy (1992). "The Division of Labor, Coordination Costs, and Knowledge." The Quarterly Journal of Economics **107**(4): 1137-1160.

Chaney, T. and R. Ossa (2013). "Market size, division of labor, and firm productivity." Journal of International Economics **90**(1): 177-180.

Edwards, A. M. (1917). "Social-Economic Groups of the United States." Publications of the American Statistical Association **15**(118): 643-661.

Fenichel, A. H. (1966). The Growth and Diffusion of Power in Manufacturing, 1839-1919. Output, employment, and productivity in the United States after 1800: Conference on Research in Income and Wealth. Studies in income and wealth. Conference on Research in Income and Wealth, National Bureau of Economic Research and Economic History Association. New York, National Bureau of Economic Research; distributed by Columbia University Press**:** xiv, 660 p.

Goldin, C. and L. F. Katz (1998). "The Origins of Technology Skill Complementarity." Quarterly Journal of Economics. **113**(3): 693-732.

Gray, R. (2013). "Taking technology to task: The skill content of technological change in early twentieth century United States." Explorations in Economic History **50**(3): 351-367.

Great Britain. Ministry of Labor (1927). A Dictionary of Occupational Terms Based on the Classification of Occupations used in the Census of Population, 1921. London, HM Stationary Office.

Hounshell, D. A. (1984). From the American system to mass production, 1800-1932. Baltimore, Johns Hopkins University Press.

Hunt, W. C. and United States. Department of Labor (1897). "Workers at Gainful Occupations at the Federal Censuses of 1870, 1880, and 1890." Bulletin of the Department of Labor **11**(July): 393-433.

Katz, L. and R. A. Margo (2014). Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective. Human Capital in History: The American Record. L. P. Boustan, C. Frydman and R. A. Margo. Chicago, University of Chicago Press**:** 15-57.

Roe, J. W. (1926). English and American tool builders. New York, London,, McGraw-Hill.

Scranton, P. (1997). Endless novelty : specialty production and American industrialization, 1865-1925. Princeton, N.J., Princeton University Press.

U.S. Congress (1894). Joint Resolution Providing for an investigation relating to the effects of machinery on labor. Washington DC.

U.S. Department of Commerce. Bureau of the Census, W. C. Hunt and A. M. Edwards (1915). Index to occupations, alphabetical and classified. Washington, Govt. Print. Off.

U.S. Department of Labor. (1899). Hand and Machine Labor (Thirteenth Annual Report). Washington DC, GPO.

U.S. Department of Labor. Employment Service (1939). Dictionary of occupational titles. Washington DC, U.S. Government Printing Office**:** 2 v. (xxxvii, 1404 p.).

United States and W. C. Hunt (1914). Thirteenth census of the United States … 1910. Volume IV. Population: 1910. Occupation statistics. Washington, Govt. Print. Off.

United States and W. C. Hunt (1914). Thirteenth census of the United States: 1910. Bulletin. Population: 1910. Occupation statistics. [Washington, Govt. Print. Off.

United States. Bureau of the Census (1910). Thirteenth census of the United States. Classified index to occupations, population division. Washington, [U.S.] Govt. Print. Off.

United States. Bureau of the Census., W. C. Hunt and A. M. Edwards (1915). Index to occupations, alphabetical and classified. Washington, Govt. Print. Off.

United States. Commissioner of Labor. (1894). Hand and Machine. C. P. National archives and records (National Archives II, MD). RG 257 Records of the Bureau of Labor Statistics. Records of the Departments of Commerce and Labor, Entry 8, Department and Bureau Forms, 1885-1913, Box. No. 3, Form No. 301-500

United States. Department of Commerce. Bureau of the Census. (1950). Alphabetical Index of Occupations and Industries: 1950. Washington, Government Printing Office.

United States. Department of Commerce. Bureau of the Census. and A. M. Edwards (1940). Alphabetical Index of Occupations and Industries. Sixteenth Census. . Washington, Government Printing Office.

United States. Department of Labor. Bureau of Employment Security. (1956). Estimates of Worker Trait Requirements for 4,000 Jobs as Defined in the Dictionarfy of Occupational Titles. Washington DC, GPO.

Ure, A. (1835). The philosophy of manufactures, or, An exposition of the scientific, moral, and commercial economy of the factory system of Great Britain. London, C. Knight.