Preliminary

Do Note Cite Without Permission

‘Mechanization Takes Command’: Inanimate Power and Labor Productivity

in Late Nineteenth Century American Manufacturing

by

Jeremy Atack

Robert A. Margo

Paul W. Rhode

July 2019

Professor Emeritus, Vanderbilt University and NBER; Professor of Economics, Boston University, and NBER; Professor of Economics, University of Michigan, and NBER. Prepared for the Economic History Association meetings in Atlanta, GA, September 2019. We acknowledge helpful comments from workshop participants at the University of Michigan; NYU-Stern; Yale University; and the Economic History @ UdeSA Conference, Buenos Aires, Argentina.

Introduction

During the nineteenth century the manufacturing sector in the United States underwent substantial growth and development, manifested by increases in the share of the labor force in the sector, growing capital intensity, and higher labor productivity (Atack, Bateman and Margo 2005; Broadberry 1998). More fundamentally, however, the entire character of manufacturing changed. At the start of the century, almost all manufacturing took place in "artisan shops," where the proprietor, perhaps with an assistant or two, fashioned goods entirely by hand using simple tools. Nothing in the production process was "automated" – there were no inanimately-powered machines that mimicked human actions.[1] However, as the century progressed, manufacturing shifted away from "hand labor" in the artisan shop to "machine labor" in mechanized factories. Specialized machines were invented that were able to duplicate certain human actions and thus displace some human labor in the production process. At the same time, these machines created a demand for labor that "operated" them. Moreover, the machines did not power themselves or use human muscle – rather, they were driven largely by inanimate sources of power, first water and later by steam (and eventually by electricity). Mechanization was closely associated with an increased division of labor -- the factory was much larger than the artisan shop, and workers specialized in tasks, unlike in the artisan shop, where specialization was limited because the number of workers was almost always far less than the number of steps involved in making a product.

[1] Throughout the paper we are equating "mechanization" with the use of steam or water power (or equivalently, that inanimate power refers to steam or water). We are aware that there were machines in the nineteenth century that used wind power or were driven by horses, or used stored energy (spring driven clocks) but these examples are minor in comparison with steam or water. Further, by the end of the century, it was steam (and eventually electricity) that was the dominant form of inanimate power.

In this paper we use a unique, extraordinarily detailed study from the late nineteenth century – the US Commissioner of Labor's 1899 Hand and Machine Labor Study (HMLS) – to measure the impact of inanimate power on labor productivity in manufacturing (United States. Department of Labor. 1899). Unprecedented for its time – and even for ours – agents for the Commissioner collected information at the production task – "operation", as the study called it – level. This information included whether or not the operation used inanimate power, the tool or machine employed, and crucially for us, the amount of time that it took to complete the task. The Commissioner selected 600+ highly specific products for study and, for each, data were collected from an establishment making the product using old-fashioned methods – hand labor – and another using "machine labor". The Commissioner's agents took great care to match the data across the two production "units" as they were called, such that it was possible to calculate, exactly how much time was "saved" by machine production in any particular operation; or, overall, for any particular good.

However, while the data have been used anecdotally to assess the productivity advantages of machine labor for a particular operation or good – they were far beyond the capabilities of statisticians at the time to use them to measure the productivity advantage on average or the role of inanimate power in generating productivity gains. As Carroll D. Wright, director of the study, put it in a magazine article in 1900, "[t]his report answers in a measure the many demands for information … but no aggregation can be made because it is impossible to carry out calculations through the innumerable ramifications of production under hand and machine methods … although such a summary would be of the greatest possible value in the study of the question of machinery." (Wright 1900)

More than a century after its original publication, advances in computing have made it possible to provide the kind of summary that Wright and his team were unable to deliver at the time. We use a digitized version of the HMLS to explore two questions. First, on average, how much more productive was machine labor than hand labor, holding constant the "content" (that is, the activity) of production tasks (insofar as this is possible)? The answer to this question turns out to be an order of magnitude: 6-7 times. In machine production, a task could be performed 6 or 7 times as often in the time it took to perform the corresponding hand labor task. With this level of productivity gain, it is not surprising that the "machine age" led to vast increases in the quantity of manufactured goods produced, and gains in consumer welfare (though we lack price data on the specific goods produced). This question, we reiterate, was not answered in the original study, because it was beyond the computational abilities of statisticians at the time given the complexity of the data.

Second, we ask how much of the productivity gain can be attributed to inanimately-powered mechanization – the use of steam or water power – *per se*. We approach this question first by estimating an OLS regression in which the dependent variable is the log difference of the amount of labor time between the machine and hand labor operation and in which the relevant independent variable is the difference in mechanization – the machine task used inanimate power, but the hand task did not. Our base OLS specification shows a large, statistically and economically significant effect of mechanization on productivity – the use of inanimate power substantially shortened the amount of time it took to complete an operation, other factors held constant. Measured at the sample means, the size of the effect implies that mechanization accounts for roughly a quarter of the higher productivity of machine labor, depending on the regression specification.

Although the HMLS has key advantages in answering our second question compared with alternative sources, we cannot rule out, *a priori*, that there is omitted variable bias. Here, the identification problem is more difficult than usual because an instrumental variable at the operation level is necessary. Our proposed solution makes use of textual information in the HMLS describing the production operations. From this text, we extract the first occurrences of all "gerunds" in hand labor operations. A gerund is an English verb to which "-ing" is appended and which, in context, functions as a noun. All told, there are 22,000+ gerunds recorded in the study, associated with the approximately 20,000 separate hand and machine operations reported. Six hundred and twenty-six gerunds are unique; these we assigned ratings regarding their technical feasibility to be mechanized by the late nineteenth century, based on the economy-wide state of knowledge in science and engineering, as described by leading historians of technology. Loosely speaking, some human activities – for example, those involving broad, repetitive movements of the arm – were understood well enough to be replicated by a machine, but other types of activities – especially those requiring fine adjustments, depending on touch, or ones that required mental judgement in the moment – were impossible to automate (Giedion 1948; Hounshell 1984). Our ordering has four ranks, from the lowest to highest feasibility. We then collapse the rank order into a 0 -1 dummy, in which "one" represents the highest rank. For each occurrence of a gerund, we compute the value of the dummy, and from this we construct our instrument.[2]

The first stage turns out to be strong, and the instrument positively and significantly predicts if a hand operation is s mechanized under machine labor. The 2SLS coefficients are

[2] We focus on the gerunds as a means of creating tractable data out of highly complex natural language. Efforts to reduce the dimensionality is a standard feature of modern analysis of "Text as Data."

positive, highly significant, and larger than the corresponding OLS coefficients, although the degree of downward bias in the OLS is not large. We also conduct robustness checks, including comparisons with the nineteenth century Census of Manufacturing (CoM) data that largely validate our substantive findings. In the end, we conclude that mechanization explains, in a causal sense, perhaps a third of the productivity gains of machine labor over hand labor, in those operations that overlapped between the two manufacturing regimes.

Economic historians have long been interested in the factors behind the diffusion of water and steam power in nineteenth century manufacturing (Atack, Bateman and Weiss 1980; Hunter 1979; Hunter 1985; Temin 1966). However, there has been relatively little previous work attempting to measure the effects of mechanization on labor productivity in manufacturing. The closest paper to this one is by Atack, Bateman, and Margo (2008) and uses the Atack-Bateman-Weiss establishment-level samples from the 1850-1880 CoM to study the impact of steam and water power on value added per worker. Like this paper, they show positive and significant effects of inanimate power on labor productivity, and the explanatory power of changes over time in power use in accounting for labor productivity growth is similar to that found here, despite significant differences in the nature of the data.

The upshot of this paper is that mechanization was, indeed, a quantitatively important factor behind the rise of American manufacturing in the nineteenth century. However, by itself, mechanization cannot account for the majority of the gains in labor productivity. Our concluding section discusses other sources of productivity gains, among which the most important may have been division of labor.

Inanimate Power and Labor Productivity in Nineteenth Century American Manufacturing

Economic historians have long studied the sources of labor productivity growth in manufacturing during the so-called "First Industrial Revolution." These sources include pure total factor productivity growth (learning-by-doing), increases in capital per worker, and an increased division of labor reflected in growth in average establishment size. Although these various sources exist independently in a theoretical sense, in reality they are interrelated and very difficult to separate. For example, new technology – such as a steam engine – is embodied in new capital goods. A new technology may also prompt changes in organization of work on the shop floor and the speed with which a particular part is produced, leading to changes in the scale of production (see, for example, Devine 1983; Hounshell 1984). Indeed, aspects of the HMLS resemble the time-and-motion studies that would become associated with Frederick Winslow Taylor (Taylor 1911a; Taylor 1911b).

Much of the previous literature on the diffusion of inanimate power in manufacturing approaches the issue from the perspective of the dual – that is, from the cost side (Atack 1979; Atack, Bateman and Weiss 1980; Temin 1966). The idea, very simple, is that reductions in the real cost of inanimate power prompted establishments to substitute it for hand or animal power. This shift was initially towards water power and, later, to steam power. The shift to steam was especially important. Steam was inherently more reliable than water, because it was not so dependent on nature; where "nature" refers to the weather and topography. By contrast, steam was "footloose"- a steam-powered establishment could locate anywhere fuel was deliverable (Kim 2005). Production, for example, could be located in cities, near thicker labor and product markets. Steam power was also scalable – an establishment could add more steam engines of a given size or it could pair a larger size engine with multiple boilers – whereas the marginal user of water power eventually rose steeply due to the difficulties and cost of expanding capacity at any given site (Hunter 1985). Indeed, many water-powered establishments installed steam engines as their power demands grew.

The evidence for the United States suggests that the pace of diffusion of inanimate power was steady over the nineteenth century, especially if one adjusts for the size of establishments in terms of the number of workers. According to Atack, Bateman, and Margo (Atack, Bateman and Margo 2008, p. 189) around 44 percent of workers in manufacturing were employed in establishments that used steam or water power in 1850. However, even though the steam engine had been invented long before, the vast majority (80 percent) of inanimately-powered manufacturing establishments used water power in 1850. The shift towards steam accelerated during the second half of the nineteenth century, particularly in larger establishments as the cost of steam power plummeted (Atack 1979; Atack, Bateman and Margo 2008). A similar slow diffusion is also present in British manufacturing. According to Crafts (2004), whose estimates derive from a growth accounting framework, it was not until well after mid-century that steam power had much of any effect on labor productivity growth in the UK; even then, the effect was very modest—a finding that is not necessarily inconsistent with ours of a more substantial impact at the establishment level and especially at the level of production operations.

The Hand and Machine Labor Study

In 1894, Congress requested that the US Commissioner of Labor investigate the “hand” and “machine” methods in commodity production (United States. Congress 1894). Although the title of the resultant study was “Hand and Machine Labor,” Commissioner of Labor Carroll D.

Wright cautioned in his introductory remarks that the words were not used in their strictest sense, but rather to characterize two different methods of production. "Machines" were used in "hand" production although these were usually simple hand tools—saws, hammers, chisels, files, knitting needles, screwdrivers and the like—what he called "the primitive method of production which was in vogue before the general use of automatic or power machines." Similarly, some tasks in machine production continued to be performed by hand using these same simple tools, including adjusting the machinery. For Wright, however, a crucial distinction was that, in machine production, "every workman has his particular work to perform, generally but a very small portion of that which goes to the completion of the article" – that is, division of labor was central to the production process (United States. Department of Labor. 1899, p. 11). The report took the Commissioner of Labor more than five years to assemble and was published in the form of a complex set of tables prefaced by some introductory text and discussion, spread across two volumes and almost 1,600 pages. We have digitized the data tables and converted them into a format that makes them amenable to standard econometric analysis

The tables describe in great detail each step in the production process for paired establishments using these different production methods to make a highly specific and carefully defined product.[3] In composing these tables, trained agents collected data were through direct observation or from written records, usually for two establishments using hand methods and two using machine methods. The data were then scrutinized and compared for inconsistencies,

[3] For example, "men's medium grade, calf, welt, lace shoes, single soles, soft box toes" (Unit 71) where 7 adjectives are used to describe the product. Some descriptions are simpler: "cane-seat maple chairs, Grecian pattern" (unit 326) or "No 7 horseshoe nails, 125 per pound" (unit 472). In just a few cases, there were minor variations in the product description between hand and machine-made products, for example, Unit 135 "Velvet carpet." In the handmade carpet, there were 780 worsted ends, 522 linen ends, and 790 cotton ends whereas the machine-made product had 216 worsted ends, 432 cotton ends, and 648 jute ends United States. Department of Labor. *Thirteenth Annual Report of the Commissioner of Labor. 1898. Hand and Machine Labor*. 1899..

missing data, and inaccuracies, following up when necessary to resolve ambiguities, before selecting the "the better and more complete one" of each mode of production of the product for presentation (United States. Department of Labor. 1899, v1, p. 13). Where necessary, production was scaled to industry norms by adjusting the time (and thus the cost) spent on tasks by the appropriate factor, keeping the number of workers unchanged. We elaborate on this below.

For machine production, the vast majority of the observations pertain to activities conducted in the mid-to-late 1890s (1894-98). For a few products, the HML staff were unable to find matching hand production from the same year that occurred nearby, presumably because the relevant establishments were no longer in existence. In such cases, the agents assiduously sought out historical records or, in 13 instances, located hand production establishments overseas that they deemed similar to those that no longer survived in the United States. All machine production data, however, was taken from US establishments.

In Part One, the following was reported for matched pair of establishments producing a highly specific product: an industry classification, an exact description of the product, the standardized quantity of that product, the year in which the production under each method took place, the number of separate tasks of production, the number of different workers employed, and the total number of hours of work to produce the given quantity, the total labor costs, and the average daily hours of operation of the unit. In Part Two, the following information was reported for each mode of producing the product: a brief description of the operation in the order in which it was performed; a list of capital goods or machines used in the operation; the type of motive power if used; the number of workers assigned to that operation; worker characteristics, such as their gender and occupation title; the hours of work by each employee engaged in the task; and the labor cost of each employee engaged in the operation along with any miscellaneous comments.

For example, the machine production of key-wound watch movements (unit 204) was broken down into 881 operations. These were listed in production order under both hand and machine production and the tasks themselves were linked across the two methods from machine production back to hand production, sub-dividing machine tasks where and if necessary and creating composite hand tasks when there was no analog between the two. We elaborate upon this below. In the hand production method for key-wound watch movements, for example, 347 tasks are identified. Most importantly, we know the *length of time* it took for each step to be completed and *the source of power* employed.

Overall, there are 672 paired units in the HML study: 27 in agriculture, 10 in mining and quarrying, and 9 in transportation, leaving 626 paired units producing manufactures. Moreover, we drop the 15 units where the hand production establishment was located outside the United States, focusing therefore on 611 domestic matched pairs as a basic sample.

Before turning to our findings, we highlight four limitations of the Hand and Machine Labor data. First, although a wide range of goods and industries are covered, the establishments that were included are in no sense a random sample either within or across industries. Second, no information was collected on output prices, revenues, or costs, except those pertaining to the labor involved directly in the production of the product (and its supervision). Consequently, any analysis of productivity, including ours, must rely on the measure provided by the study—the amount of time that it took to complete an operation. Third, while the agents recorded additional information on the survey form that would have been very useful to have for some analyses—for example, the names of the individual workers, and the address of the establishment—this information was not included in the published study. Moreover, as far as we can determine, the completed survey forms have not survived and so this additional information has been lost. Finally, the study reported the labor requirements for a standardized scale of production, which enhances comparability. But the number of workers employed and the organization of work may not reflect how producers, especially hand producers, would have operated at that specific (usually larger) scale under more realistic time-cost considerations.

Operation Blocks and Block Links

The fundamental objective of the HMLS was to enable comparisons to be made between hand and machine labor, either at the product (equivalently, what the HMLS termed the “unit”) level or at the operation level (that is, production activities). To do this, the HML staff took a series of steps. First, they ordered the machine operations from start to finish, giving each a consecutive number, which was listed in a column labelled “Operation number”. For example, the production of a 14-tooth steel garden rake by machine labor (unit #30, volume 2, p. 480) took 16 operations; the first, operation #1, was “cutting iron into pieces” using shears while the last, operation #16, was “inspecting rakes and overseeing the establishment”. The staff did the same for the hand labor operations, except that in the “Operation number” column, there could be letters, numbers, or combinations of both where the numbers matched up with those from machine production. It is these entries in the hand labor “Operation number” column that provide a “crosswalk” between the hand and machine operations. To understand this better, it is useful to introduce the concepts of an operation “block” and operation “block link”, although the HMLS did not actually use these terms. An operation block is a collection of operations of size H (for hand labor) or M (for machine labor), where H and M are integers equal to or greater than one. A “block link” is a mapping between the hand and machine blocks. Some hand operations could not be matched to any machine operations because the operations were archaic or “old” – that is, was no longer performed under machine labor. We refer to these as 1:0 block links. Analogously, some machine blocks could not be matched to any hand blocks because the machine operations were novel – that is, were not performed under machine labor. We call these 0:1 block links. All other block links are designated H:M, where H and M are both equal to or greater than one, although not necessarily equal to each other.

The H:M block links are key to our analysis, because by differencing between hand and machine labor within these links we eliminate a “block-product” fixed effect (see below). That is, when we difference within such links, we are making comparisons between hand and machine labor such that the underlying production activities are held constant so that the product at the conclusion of that stage or activity was the same, except for the possibility that the machine operations might be mechanized to a greater extent than the hand operations; hence, by differencing, we get rid of the fixed effect. Note, however, that, we cannot make the same argument for the “old” (1:0) or “new” (0:1) links because, by definition, operation content for these links did not overlap.[4] As such, the 1:0 and 0:1 links are excluded from our regression analysis.

Table 1 shows the distribution of 1:0, 0:1, and most common H:M links from the perspective of hand labor (Panel A) or machine labor (Panel B). In each panel, we also report the mean fraction using steam, water, or “mechanized”, (used steam and/or water). The overall means are equally weighted averages across the producing units. For hand labor, there were a total of 6,208 operation blocks. Of these, 431, or just 6.9 percent, were 1:0 links, that is, hand

[4] We could aggregate the 1:0 and 0:1 blocks to the unit level, and then difference between labor types. However, these differences would be perfectly correlated with the product fixed effects that we include in the regression and thus would be uninformative.

operations that were not practiced under machine labor. The HML staff recognized that, in an ideal study design, literally none of the hand labor operations would have used inanimate power. Despite strenuous efforts on the part of the staff, they were not able to achieve this ideal – but as can be seen, the fraction of hand operations that made use of steam or water was extremely low.

For machine labor, there were a total of 10,017 blocks – significantly more than the number of hand labor blocks. This discrepancy reflects, in large part, the relatively large number of machine operations – 4,240 or 42.3 percent – that were 0:1 links that is, had no counterpart under hand labor (such as tending to the power source). On average, 56 percent of the machine blocks that overlapped with the hand blocks (labelled "regression sample" in Table 1) were mechanized, the vast majority of which (53 percent) used steam power alone. Interestingly, the degree of mechanization was greater for those blocks in which some subdivision of operations took place in the transition from hand to machine labor (1:M blocks) or, alternatively, consolidation (H:1). In our main empirical analysis, we will pool the data across the relevant block links but include block link dummies as a control. Later, in our robustness checks, we will allow the effects of power to differ between the 1:1 and non-1:1 block links.

Empirical Analysis

We use operation block level data, as described in the previous section, as the basis for our analysis of the effects of mechanization on labor productivity. As the HML staff clearly understood, if it was possible to change something in the production process – for example, use steam-powered machinery – to complete the activities in an operation block more quickly than before – productivity would be increased. We derive our base regression specification from the following equation:

$$\text{Ln } T(i,j,k) = \alpha(i,k) + \beta(j,k) + \gamma(i,j) + \lambda^*(\text{Steam} = 1|i,j,k) + \theta^*(\text{Water} = 1|i,j,k) + \varepsilon(i,j,k)$$

The index i refers to the block; the index j, to the type of labor (j = hand or machine); and the index k, to the specific product, what the HML staff called the "unit". Ln T(i,j,k) is the log of the amount of time that it takes to complete the operation block for labor type j in unit k. The parameter α is a block product fixed effect; that is, it is indexed for block i and unit k but not for labor type j. In making α dependent on i and k but not j, we are assuming that, while some blocks might take proportionately longer than others for a given product, these relative differences are the same under both machine and hand labor. The parameter β is a labor type-product fixed effect; it is indexed by j and k but not by i. This allows for the possibility that machine labor was more productive in general and that the productivity gain differed across specific products. The parameter $\gamma(i,j)$ is a block-labor type fixed effect; when we difference the data (see below) this results in a set of block type fixed effects. Our main interest is in the parameters λ and θ which are the log effects of steam and water power use. If steam or water power use proportionately reduces the amount of time to complete a block, then $\lambda < 0$ and $\theta < 0$.

To estimate this regression, we need, as discussed previously, to compute differences between machine and hand labor within units for all variables measured at the block level. For the 1:1 links, we can do this simply and directly because for every machine operation there is an exact counterpart operation under hand labor. For the other block links, it is necessary to compute averages or totals at the block level before taking differences. In the case of the motive power variables on the right-hand side, this involves computing the proportion of operations in the block that use a particular power source, so we refer to these variables, for example, as "Fraction Steam" or "Fraction Water" as appropriate.

After differencing, we have:

$$\Delta \ln T = \Delta\beta + \Delta \gamma + \lambda^* \Delta (\text{Steam} = 1) + \theta^* \Delta (\text{Water} = 1) + \Delta \varepsilon$$

For ease of reading, we suppress the indexes above but keep in mind that the unit of observation is the task. The dependent variable is the difference between machine and hand labor in the log of the amount of time that it took to complete a task, ($\Delta \ln T$). The right-hand side variables are product fixed effects ($\Delta\beta$), block-link type fixed effects ($\Delta\gamma$), the differences in the steam and water power dummies between machine and hand labor at the task level, and the difference in the error terms ($\Delta \varepsilon$).

Our first approach is to estimate the regression using OLS. To interpret the power coefficients as casual effects, we must assume that changes in inanimate power use at the operation block level were uncorrelated with the error term, $\Delta \varepsilon$. Below we relax this assumption when we develop an instrumental variable strategy for mechanization.

Column 1 of Panel A in Table 2 reports the equation as specified above along with those from alternative right hand side specifications. Instead of reporting on the effects of steam and water separately, we can collapse the power variables, Δ Fraction Steam and Δ Fraction Water, into a single variable, Δ Fraction Mechanization. This variant, reported in column 3 of Panel A, is used as our baseline when we turn to the IV analysis below. Substantively, however, the change to a single power variable has very little impact because the vast majority of changes in power in the HLMS involve the adoption of steam. One can also include the log of the time to complete the hand labor block as a right-hand side variable. This specification is analogous to including a lagged dependent variable in a panel regression. As such, we anticipate that the coefficient of this "lagged" variable will likely be negative, and that the magnitude of the power coefficients will be reduced. The effects of including this variable are reported in columns 2 and 4 of Panel A.

The power coefficients are uniformly negative and highly significant, indicating that mechanization reduced the amount of time to complete the operation block and, therefore, raised labor productivity. The marginal effect of steam, in this regard, was much larger than water, indicating the superiority of steam as a source of power. Collapsing the power dummies into a single mechanization variable only slightly reduces the magnitude of the power coefficient as one might expect because almost all of the mechanization documented in the HMLS involved steam power. As anticipated, including the log of the time spent in the hand task reduces the magnitude of the power coefficients, although they remain relatively large and statistically significant.

In Panel B of Table 2 we use the OLS coefficients and the mean values to compute the "percent explained" of the average productivity difference between hand and machine labor by mechanization. As shown in Panel B, this average productivity gain, in log terms, was -1.90. If we take the exponent, we get exp [-1.90] = 0.15, or the average machine block took just 15 percent of the time to complete than its counterpart under hand labor – that is, machine labor was 6-7 times as productive as hand labor, using the HMLS metric.[5]

How much of this impressive productivity gain can be attributed to mechanization? The answer is shown in the bottom two rows of Panel B. In the penultimate row, we show the predicted change in the dependent variable based on the regression coefficients applied to the mean values of the mechanization variables. In the last row, we divide the predicted change by the mean value of the dependent variable. According to this calculation, between 20 and 28 percent of the gains in productivity can be attributed to mechanization. Thus, mechanization clearly was a quantitatively important factor behind the greater productivity of machine labor but it does not account for the majority of the productivity difference, let alone all of it. However, this assumes that the OLS coefficients are true causal estimates and not, in particular, biased downwards through omitted variable bias. We now turn to investigating this further.

Instrumental Variable Analysis: Gerunds

We develop in this section an instrumental variable to measure the effects of mechanization in the HMLS. Aside from satisfying the standard relevance and exogeneity conditions, we need to measure any instrumental variable for our application at the operation block level. This makes our problem especially difficult; even if we had more information on the establishments included in the HMLS, the variation would likely be at the unit level, not the operation block level.

[5] This slightly overstates the productivity advantage of inanimate power because the HMLS recorded the labor time of workers "furnishing power" (e.g. operating the steam engine) as a new operation (0:1 block link). However, these activities took up on a tiny fraction of the total time of new activities (the mean time share was 2.6 percent), indicting that any overstatement of the productivity gains is minimal.

To make progress on solving this problem, we make use of the written descriptions of operations in the HMLS which was recorded as “Work Done.”[6] These contain one or more gerunds, describing activities. We have extracted the first gerund listed, called the “principal gerund” and use these to construct our instrument. To fix ideas, a gerund is an English verb to which “-ing” has been appended and functions as a noun in grammatical context. For example, consider the gerund “reading” derived from the verb “read”. In the sentence, “I enjoy reading”, “reading” is a gerund, a noun that describes an action. They are active. All gerunds end in “ing” but not all words ending in “ing” are gerunds. For example, in the sentence “I am reading a book”, “reading” is not a gerund, but rather the present participle.

Our extraction of gerunds from the text of “Work done” produced a total of 22,734 gerunds, of which 626 are unique. Some gerunds are fairly specific, like “cutting,” “drilling,” and “fastening.” For example, “cutting” appears 14 times in the 72 separate tasks or groups of tasks listed in the hand production of men’s medium grade shoes (Unit 71). The same term appears in 22 of the 173 distinct operations for the machine production of the same product. Indeed, cutting was the single most common gerund and was used in describing 2,409 tasks. It involved both organic materials like leather, paper and textiles as well as metals. Other gerunds are relatively vague, like “assembling,” “making” or “inspecting.” In the majority of operations, a single gerund was used to describe the activity undertaken, although there were as many as five gerunds in describing a task (i.e. a row in one of the tables in Part Two of the HMLS). Some gerunds are very similar to one another such as “counterboring” and “countersinking,”

[6] Our extraction and analysis of the text data on gerunds broadly follows the protocol outlined in Gentzkow, Matthew, Brian T. Kelly, and Matt Taddy. "Text as Data." *Journal of Economic Literature* (forthcoming): (see also NBER Working Paper No. 23276, March. Cambridge MA: NBER)..

"cleaning" and "cleansing," and "joining" and "jointing".[7] Moreover, sometimes the gerund was prefixed with an adjective such as "rough," "final" or "finish." Despite similarities and modifiers, however, each unique gerund is treated as a different activity.[8]

Some activities—like cutting, drilling, boring and turning—were mechanized quite early using successful devices like Wilkinson's boring machine (1775) or Blanchard's lathe (1818). Other activities still defy mechanization as they involve too much idiosyncratic decision-making or complexity as in "overseeing," "examining," "finishing," "inspecting," "assembling" and "repairing." Certain activities require attention and judgment—like assisting, examining, fitting, gauging, tending, and beaming warp—which refers to the tensioning of warps on a weaving frame which have to be tensioned "just-so"—tightly but not so tightly that they will break if subjected to a little extra pressure from say the weft and that will vary with humidity, the type and possibly the batch of thread used for the warp.

We have used this insight to generate a four-way classification for each gerund based upon our judgment about how easily these activities could be mechanized based upon existing technology by the late nineteenth century, ranking from high to low. For example, activities like annealing, boring, drilling, or turning are viewed as involving purely technical issues that had already been solved and were ranked high. Some of these issues were mechanical, as in boring/drilling and involved experiential or experimental decisions like determining the "right"

[7] Some gerunds are very similar and we would be hard pressed to distinguish between them in terms of the activity—for example, cleaning and cleansing. Similarly, while a joint is the place where two or more pieces are joined together and there are "jointing machines" (which, in woodworking, could include joints made by rabbits, dadoes, mortises and tenons and so on). However, there is a small distinction between counterboring and countersinking in that a bore is straight-sided whereas the sink is tapered. In each case, though the objective is to let a screw or bolt sit at or below the surface. Despite similarities, each gerund was classified separately.

[8] We also paid special attention to certain gerunds—for example, "striping" and "stripping," which look very similar to one another but are not—the one we think of as adding material; the other, of removing material. In these cases, the original text was carefully checked to verify that they were not a product of mistaken transcription or a result of type-setting or spelling errors in the assembly of the text. We concluded that they were not.

metal for the bit and best speed for the parts being worked. Activities like annealing and tempering, however, once the metallurgical issues were resolved, were simply time-dependent. Other tasks, like grinding and trimming, could be mechanized but may have required some product redesign so the item would register properly or required some minimal active oversight of the operation. These activities were ranked medium high. Activities like "fitting," "forging," "gluing" and "cementing" required (and often still require) more idiosyncratic judgement and were assigned a medium-low ranking. Fit, for example, is often determined by how much resistance is involved in putting the parts together and how much "wiggle" remains once they are in place, issues that are difficult to teach and measure. Lastly, there were those activities, like "assisting," "overseeing," "examining," "finishing" and "assembling" where human judgement was essential. These have a low potential for mechanization or automation.

This four-way classification was then used to construct a dummy variable, MECHABLE, assumed a value of 1 if and only if the potential for mechanization had been judged "high," otherwise zero. Because we have extracted the complete list of gerunds, for any given hand labor block we can compute the variable, "Fraction MECHABLE," which is the proportion of gerunds for which MECHABLE equals one. We are then in a position to estimate the first stage, which is a regression of Δ (Fraction Mechanized) on Fraction MECHABLE. This regression also includes product and block type fixed effects.

As shown in Table 3, the first stage works well – the coefficient is positive and highly significant, indicating that if a gerund had strong potential to be mechanized on the basis of economy wide scientific and engineering knowledge, the probability that the task was, in fact, mechanized under machine labor was higher. The 2SLS coefficients, also shown in Table 3, are negative, highly significant, and larger in magnitude than the OLS coefficients, indicating a downward OLS bias.[9] One possible explanation for the downward bias is unobserved worker quality – specifically, that the machine worker was less able than the hand worker, doing the same operation. Correcting for this bias would increase the magnitude of the power coefficient, which is what we observe.

We use the 2SLS coefficient in conjunction with the mean values of the dependent variable and the change in fraction mechanized to compute the percent explained of the change in labor productivity. Because the 2SLS coefficient is larger in magnitude, we now attribute more of the change in labor productivity to mechanization – depending on the regression specification, between 31 and 33 percent. As before, therefore, mechanization was clearly important to raising labor productivity in manufacturing but on its own it does not account for the majority of the gap in productivity between hand and machine labor.

Robustness Checks

We performed several robustness checks of our base OLS and IV analyses. Panels A and B of Table 4 summarize the results of the robustness checks.

Our base analyses treat each block link observation equally. An alternative is to weight observations by the total number of operations in the block – for example, if an observation is a 1:M block, the weight is 1 + M. As can be seen in Table 4, weighting in this manner very slightly reduces the coefficient magnitudes, but these remain highly negative and significant.

[9] We also estimated the 2SLS coefficients by modifying the IV slightly, such that MECHABLE = 1 if the potential for mechanization was judged to be medium high (an index of 3) or high (an index of 4). The 2SLS coefficients were still negative and highly significant and only slightly smaller in absolute value than as shown in Table 3 (for example, -1.10 instead of -1.25 (row 6, column 2).

In Tables 2 and 3 we pooled the data across block link types to estimate the OLS and IV regressions. Some pooling is necessary because the sample sizes are too small to perform the estimation separately by block type. However, we can split the analysis between the 1:1 and non-1:1 block link types, and see if the relevant coefficients differ across the types. As can be seen in Table 4, we find some evidence that the power coefficients differed in size between the 1:1 and non-1:1 links, but any such difference largely disappears once we control for the log of time spent in the hand labor block as a right hand side variable.

We noticed that a relatively large fraction (14 percent) of observations in the regression sample come from the small number (five) of units that made watches or clocks (units 201-205). Watches and clocks were part of the famed "American System" of manufactures, which emphasized mechanization, division of labor, and widespread use of interchangeable parts (Hounshell 1984, pp. 51-61). It is reasonable, therefore, to wonder if our base results primarily reflect the impact of this industry. However, dropping watches and clocks from the sample, if anything, strengthens the effects of mechanization.

Comparison with Census of Manufactures

Atack, Bateman and Margo (2008) use establishment level data from the 1850-1880 censuses to study the impact of steam and water power on labor productivity. We estimate a variant of their base regression specification using these same data, and compare those results with those from the HMLS.

The two data sources are very different – value added per worker in the CoM is nothing like the measure of the labor productivity in the HMLS data – time spent in an operation or in fashioning a specific good from start to finish – although the measures are surely correlated. In the HMLS analysis, we are trying to explain, to a first approximation, contemporaneous differences in productivity between two types of labor regime whereas in the CoM analysis we are explaining overall changes in labor productivity over time. We can, as in the HMLS, measure mechanization by steam and water power separately, and we can also do so with dummy variables, which is effectively what we are doing in our analysis of the HMLS operations level data.

The dependent variable in our analysis of the CoM data is the same as that used in Atack, Bateman, and Margo (2008), the log of value added per worker. We pool the data across the four census years. The main independent variables are dummies for use of steam power and for water power. We also include flexible controls for establishment size; the fraction of workers who were female; a dummy for urban status; fixed effects for year, state, and 3-digit SIC code; and year-region and year-SIC interactions. Observations are weighted by the number of workers in the establishment, again following Atack, Bateman and Margo (2008).

Table 5 shows the coefficients of the steam and water dummies, and the results of a percent explained calculation analogous to those in Tables 2-4. The power coefficients are positive and statistically significant; consistent with the findings reported in the original Atack, Bateman, and Margo (2008) paper, the marginal impact of steam power is larger than that of water power. If we use the coefficients in conjunction with the mean values of the power dummies in 1850 and 1880 to predict the change in labor productivity over the period, the predicted change is 0.076 log points, compared with an actual change in log real value added of 0.244 log points. Thus, shifts in power use, as revealed in the CoM data, account for approximately 31 percent of the rise in labor productivity in US manufacturing between 1850 and 1880.

Our estimates using the CoM are OLS, and therefore subject to the usual critiques of omitted variable bias and endogeneity; it is certainly possible that an IV analysis of the census data (assuming we had plausible instruments) would show downward bias in the OLS effects, similar to what we observe in the HMLS data. That said, both the CoM and HML data agree that mechanization was a significant factor in raising labor productivity, but by itself does not account for the majority of the productivity gap we are interested in explaining; and it seems doubtful that this substantive conclusion change if we had a true causal analysis of the CoM data.

Discussion and Concluding Remarks

Over the course of the nineteenth century the US manufacturing sector ascended, achieving levels of labor productivity that were substantially higher than in Great Britain, the world's first industrializing nation. American manufacturing ascendancy went hand in hand with the shift of production and factor inputs from the "hand labor" of the non-mechanized artisan shop to the mechanized factory.

It has long been an article of faith among economic historians that the use of inanimate power was a critical factor in the secular increase in labor productivity in American manufacturing. We have used a unique and extraordinarily detailed late nineteenth century data set, the Department of Labor's Hand and Machine Labor Study, to analyze the effects of mechanization. We present OLS and IV estimates of the "treatment effects" of mechanization at the production operation level. Using our IV estimates, roughly a third of the higher degree of labor productivity in machine production is attributed to mechanization per se. This is similar in size to that based on a re-analysis of the 1850-1880 census of manufactures data used by Atack, Bateman, and Margo (2008)and is important because the two data sources – the HMLS and CoM – are fundamentally different.

On the one hand, our findings certainly support modern scholars, both economic historians and economists, who argue that the steam engine was the first great "general purpose technology" (GPT), and that historical automation, particularly in its productivity effects, bears many similarities to automation today. On the other hand, it is clear that mechanization, by itself, cannot account for the majority, let alone the entirety of the difference in labor productivity, between hand and machine labor in the HMLS; and the same is true for the growth in labor productivity evident in the CoM from 1850 to 1880.

If mechanization alone does not explain the vastly higher level of machine labor productivity in the HMLS, what other factors might? Arguably, one such factor is division of labor. Ever since Adam Smith, economists have known that division of labor, by itself, could raise labor productivity, as workers are allocated to production tasks based on comparative advantage and saving on the "costs" of shifting tools between tasks.

In order to incorporate division of labor into the analysis in this paper and run a "horse race" between it and mechanization, we would need to be able to measure it at the level of the operation block. This would be feasible if we could recover the complete assignment of the individual workers to production operations in the HMLS. In turn, we could do this if we knew the names of the workers or had some other way of uniquely identifying individuals.   The names were included in the original survey forms but not published due to privacy concerns. Consequently, we cannot include measures of the division of labor in regressions of the HMLS data at the operation block level. Nor did the census of manufactures collect data on the division of labor.

However, while we cannot measure the division of labor at the operation block level in the HMLS, we can measure it at the production unit level. In an earlier paper (Atack, Margo and Rhode 2017) we presented OLS estimates of an equation similar to our base specification, in which the dependent variable was the logarithm of the time taken to produce the product in its entirety, not just a specific operation block. In that regression, there were two observations per product, one pertaining to hand labor and the other to machine labor. In our base specification in the earlier paper, we included a dummy variable for hand labor, product fixed effects, and a log linear term in the number of workers. The coefficient of the latter is negative, implying that larger establishments (in terms of the number of workers) were more productive, consistent with economies of scale.

At the production unit level, we are able to construct two measures of the division of labor, the logarithm of the total number of operations performed in making the product, and the proportion of operations performed by the average worker. When we include these variables in the regression, their coefficients imply that an increase in the division of labor results is associated with higher labor productivity. Moreover, including these variables completely eliminates the negative coefficient on the number of workers, suggesting that division of labor was the reason behind the economies of scale. Including the division of labor variables reduces the coefficient on the hand labor dummy from 1.45 to 1.10, implying that division of labor accounts for approximately 30 percent of the higher productivity of machine labor – not very different, in terms of magnitude, from mechanization.

The analysis in Atack, Bateman, and Margo (2008) uses OLS. We do not have, at present, any instruments for Atack, Bateman, and Margo's unit level measure. However, it seems unlikely that correcting for it would double the explanatory power of division of labor. This would be necessary if division of labor and mechanization were the sole factors behind labor productivity growth.   Advances in management, including more precise inventory control due to better and more reliable transportation, improvements in financial markets, and higher average levels of human capital in workers at the end of the century than at the beginning, may also have mattered, but these are not factors that are measured in the HMLS (or, for that matter,in the CoM).

In the 1890s, Congress asked the US Commissioner of Labor to investigate the effects of the transition from hand to machine labor on labor productivity, resulting in the Hand and Machine Labor Study.   The HML staff managed to put together a remarkable body of data to answer Congress's charge, but it was unable to analyze it much at the time. Some 125-odd years later, advances in computer technology have finally opened up the HMLS to statistical scrutiny, allowing it to shed light on fundamental features of America's historical industrialization.

Tables

Table 1: Block Links, Hand and Machine Labor Study

Panel A: Hand Labor

Block Link	Number of Links	Fraction Steam	Fraction Water	Fraction Mechanized
1:0	431	0.009	0.007	0.016
1:1	4,275	0.014	0.035	0.048
1:M, M > 1	897	0.029	0.009	0.038
N:M, N, M > 1	220	0.007	0.018	0.025
N:1, N> 1	385	0.008	0.139	0.146
Total	6,208	0.015	0.035	0.050
Total, Regression Sample	5,777	0.016 [0.014]	0.037 [0.037]	0.052 [0.051]

Notes to Panel A: computed from digitized version of Hand and Machine Study, see text and (United States. Department of Labor. 1899). 1:0: hand labor operations that “disappeared” (have no counterpart) in the transition to machine labor. 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. N:M: A block of N hand labor operations is mapped to a block of M machine labor operations, N and M > 1. N:1, N> 1: A block of N (>1) hand operations is mapped to a single machine labor operation. Regression sample excludes 1:0 observations by construction. Mechanized = 1 if operation used steam or water power or both. Fraction Steam, etc: is computed in two steps: (1) within each block, fraction of operations that use steam or water (2) overall fraction is equal weighted average across units. [ ] computed in two steps (1) time-weighted average across within the block (2) overall figure is equal weighted average across units

Panel B: Machine Labor

Block Links	Number of Links	Fraction Steam	Fraction Water	Fraction Mechanized
1:1	4,275	0.485	0.028	0.508
1:M, M> 1	897	0.621	0.037	0.625
N:M, N, M > 1	220	0.694	0.033	0.724
N:1, N> 1	385	0.738	0.034	0.770
0:1	4,240	0.363	0.012	0.372
Total	10,017	0.459	0.022	0.479
Total, Regression Sample	5,777	0.531 [0.529]	0.030 [0.030]	0.557 [0.555]

Notes to Panel B: computed from digitized version of Hand and Machine Study, see text and (United States. Department of Labor. 1899). 1:0: hand labor operations that “disappeared” (have no counterpart) in the transition to machine labor. 1:1: a single hand labor operation is mapped to a single machine labor operation. 1:M: a single hand labor operation is mapped to a block of M machine operations. N:M: A block of N hand labor operations is mapped to a block of M machine labor operations, N and M both greater than 1. N:1: A block of N hand operations is mapped to a single machine labor operation. 0:1: new operations under machine labor. Regression sample excludes 0:1 observations by construction. Mechanized = 1 if operation used steam or water power or both. Fraction Steam, etc: is computed in two steps: (1) within each block, fraction of operations that use steam or water (2) overall fraction is equal weighted average across units. [ ] computed in two steps (1) time-weighted average across within the block (2) overall figure is equal weighted average across units

Table 2: The Productivity Effects of Inanimate Power in the Transition from Hand to Machine Labor: OLS Estimates and Percent Explained

Panel A: OLS Estimates, Productivity Effects of Inanimate Power

Dependent Variable	Δ Ln (Time spent in block)	Δ Ln (Time spent in block)	Δ Ln (Time spent in block)	Δ Ln (Time spent in block)
Ln (Time spent in hand labor block)		-0.38 (10.96)		-0.38 (11.20)
Δ Fraction Steam	-1.04 (19.27)	-0.76 (15.22)		
Δ Fraction Water	-0.35 (3.29)	-0.22 (2.24)		
Δ Fraction Mechanized			-0.99 (17.55)	-0.72 (14.93)
Adjusted R2	0.500	0.597	0.500	0.594

Note: sample consists of 1:1, 1:M, N:1 and N:M block links for which there was complete information on the relevant variables (N = 5,747). See Table 1 for source information. Δ: difference between machine and hand labor. All regressions include product fixed effects and dummy variables for block type (the left out dummy is 1:1). Standard errors are clustered at the product level. Absolute values of t-statistics are shown in parentheses

(Table 2 continued)
Panel B: Percent Explained of Δ Ln (Time) by Δ Power

Regression includes Ln (Time spent in hand labor block)?	No	Yes	No	Yes
Mean Value of Dependent Variable	-1.90	1.90	-1.90	-1.90
Mean Value of Δ Fraction Steam	0.52	0.52		
Mean Value of Δ Fraction Water	-0.007	-0.007		
Mean Value Δ Fraction Mechanized			0.51	0.51
Predicted Change in Mean Value of Dependent Variable	-0.54	-0.39	-0.51	-0.37
Percent Explained (Predicted Change/Mean Value of Dependent Variable) x 100 percent	28.4%	20.7%	26.6%	19.5%

Predicted change in mean value of dependent variable: computed by multiplying OLS coefficients of power use dummies (e.g. Δ (Steam = 1) by mean value of change in power use (e.g. Δ (Steam = 1). Example, column 1: -0.54 = (-1.04 x 0.52) + (- 0.35 x -0.007).

Table 3: Instrumental Variable Regressions

Dependent Variable	Δ Ln (Time spent in operation block)	Δ Ln (Time spent in operation block)
Includes Ln (Time spent in hand operation block)?	No	Yes
First Stage Coefficient	0.23 (11.27)	0.23 (11.78)
Mean Value of Instrument (Fraction MECHABLE)	0.62	0.62
2SLS:		
Δ Fraction Mechanized	-1.25 (7.10)	-1.18 (8.06)
Predicted Change in Dependent Variable	0.64	-0.60
Percent Explained	33.6%	31.7%

Instrument variable (Fraction MECHABLE) is based on gerunds in the text description of the hand operation; see text. First stage coefficient: coefficient of instrument in a regression of Δ (Mechanized = 1) with product and block link type fixed effects, column 2; or product and block type fixed effects and ln (Time spent in hand labor block), column 3. Standard errors clustered at unit level. Absolute t-statistics shown in parentheses. Percent Explained: predicted change in dependent variable using 2SLS coefficient and mean value of Δ Percent Mechanized, divided by mean value of dependent variable; see Table 2, Panel B.

Table 4: Robustness Checks

Panel A: OLS Coefficients

Robustness Check	Sample Size (number of block links)	OLS, Δ Fraction Steam	OLS, Δ Fraction Water	OLS, Δ Fraction Steam	OLS, Δ Fraction Water
Include ln (time in hand block) on right hand side?		No	No	Yes	Yes
Weighted, pooled sample	5,745	-0.99 (15.91)	-0.36 (3.07)	-0.74 (13.76)	-0.22 (2.00)
1:1 block links	4,265	-1.12 (18.73)	-0.35 (2.79)	-0.84 (15.22)	-0.27 (2.26)
Non-1:1 block links	1,482	-0.75 (5.03)	-0.41 (1.05)	-0.64 (5.18)	-0.18 (0.45)
Pooled sample, drop Watch/Clock units	4,940	-1.07 (17.81)	-0.35 (1.98)	-0.79 (14.45)	-0.24 (1.41

Weighted sample: observations are weighted by the number of underlying operations in the block link; for example, if the observation is a 1:M link, the weight is 1+ M. Watch/clock units are units 201-205.

Panel B: IV Coefficients

Robustness Check	2SLS, Δ Fraction Mechanized	Percent Explained	2SLS, Δ Fraction Mechanized	Percent Explained
Include ln (time in hand block)	No	No	Yes	Yes
Weighted, pooled sample	-1.15 (4.72)	29.9%	-1.11 (6.59)	28.7%
1:1 block links	-1.36 (8.26)	36.5%	-1.17 (7.48)	31.4%
Non-1:1 block links	-0.65 (1.01)	17.0%	-0.95 (3.01)	24.8%
Pooled sample, drop Watch/Clock unites	-1.44 (10.18)	42.0%	-1.26 (9.27)	36.8%

Table 5: Census of Manufactures Estimates and Percent Explained

Coefficient of Steam Dummy	0.242 (17.15))
Coefficient of Water Dummy	0.039 (2.13)
Adjusted R-Square	0.248
Δ Mean, 1850-1880 of:	
Ln (Real value added/worker)	0.244
Fraction Steam	0.373
Fraction Water	-0.159
Predicted Δ in ln (real value added/worker)	0.076
Percent Explained [= (row 8/row 5) x 100 percent]	31.1%

Rows 1, 2: coefficients from a regression analysis of ln (value added/worker) using the sample of establishments from the 1850-1880 censuses of manufacturing analyzed by Atack, Bateman, and Margo (2008). See text for discussion of regression specification. 1880 observations are reweighted to account for under-reporting in the original census; see Atack, Bateman, and Margo (2008). Nominal value added converted to real value added using the output price deflator in Atack, Bateman, and Margo (2008).

Bibliography

Atack, Jeremy. "Fact or Fiction? The Relative Costs of Steam and Water Power: A Simulation Approach." *Explorations in Economic History* 16, no. 4 (1979): 409-37.

Atack, Jeremy, Fred Bateman, and Robert A Margo. "Capital Deepening and the Rise of the Factory: The American Experience During the Nineteenth Century." *Economic History Review* 58, no. 3 (2005): 586-95.

Atack, Jeremy, Fred Bateman, and Robert A. Margo. "Steam Power, Establishment Size, and Labor Productivity in Nineteenth Century American Manufacturing." *Explorations in Economic History* 45, no. April (2008): 185-98.

Atack, Jeremy, Fred Bateman, and Thomas Weiss. "The Regional Diffusion and Adoption of the Steam Engine in American Manufacturing." *Journal of Economic History* 40, no. 2 (1980): 281-308.

Atack, Jeremy, Robert A Margo, and Paul W Rhode. "The Division of Labor and Economies of Scale in Late Nineteenth Century American Manufacturing." unpublished working paper No. Unpublished, 2017.

Broadberry, Stephen N. "How Did the United States and Germany Overtake Britain? A Sectoral Analysis of Comparative Productivity Levels, 1870-1990." *The Journal of Economic History* 58, no. 2 (1998): 375-407.

Crafts, Nicholas. "Steam as a General Purpose Technology: A Growth Accounting Perspective." *Economic Journal* 114 (2004): 338-51.

Devine, Warren D. "From Shafts to Wires: Historical Perspective on Electrification." *Journal of Economic History.* 43, no. 2 (1983): 347-72.

Gentzkow, Matthew, Brian T. Kelly, and Matt Taddy. "Text as Data." *Journal of Economic Literature* (forthcoming): (see also NBER Working Paper No. 23276, March. Cambridge MA: NBER).

Giedion, S. *Mechanization Takes Command : A Contribution to Anonymous History*. New York: Oxford University Press, 1948.

Hounshell, David A. *From the American System to Mass Production, 1800-1932 : The Development of Manufacturing Technology in the United States*. Baltimore: Johns Hopkins University Press, 1984.

Hunter, Louis C. *A History of Industrial Power in the United States, 1780-1930: Waterpower in the Century of the Steam Engine*. Charlottesville: University Press of Virginia. , 1979.

———. *A History of Industrial Power in the United States, 1780-1930: Steampower*. Charlottesville.: University Press of Virginia., 1985.

Kim, Sukkoo. "Industrialization and Urbanization: Did the Steam Engine Contribute to the Growth of Cities in the United States?" *Explorations in Economic History* 42 (2005): 586-98.

Taylor, Frederick Winslow. *The Principles of Scientific Management*. New York ; London: Harper & Brothers, 1911a.

———. *Shop Management*. New York and London,: Harper & Brothers, 1911b.

Temin, Peter. "Steam and Waterpower in the Early Nineteenth Century." *Journal of Economic History*. 26 (1966): 187-205.

United States. Congress, . *Joint Resolution Providing for an Investigation Relating to the Effects of Machinery on Labor*. 1894.

United States. Department of Labor. *Thirteenth Annual Report of the Commissioner of Labor. 1898. Hand and Machine Labor*. 1899.

Wright, Carroll D. "Hand and Machine Labor." *Gunton's Magazine*, March 1900, 209-17.