NBER WORKING PAPER SERIES

ECONOMIES OF SCALE IN NINETEENTH CENTURY AMERICAN MANUFACTURING REVISITED: A RESOLUTION OF THE ENTREPRENEURIAL LABOR INPUT PROBLEM

Robert A. Margo

Working Paper 19147
http://www.nber.org/papers/w19147

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2013

This is a draft of a paper to be presented at the NBER conference on "Enterprising America: Businesses, Banks, and Credit Markets in Historical Perspective" to be held at Vanderbilt University, Nashville TN, in December 2013. Comments from Jeremy Atack, Stanley Engerman, and Ian Keay are gratefully acknowledged. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Economies of Scale in Nineteenth Century American Manufacturing Revisited: A Resolution of the Entrepreneurial Labor Input Problem

Robert A. Margo

NBER Working Paper No. 19147
June 2013
JEL No. N61

ABSTRACT

In a famous paper, Kenneth Sokoloff argued that the labor input of entrepreneurs was generally not included in the count of workers in manufacturing establishments in the early censuses of manufacturing. According to Sokoloff, this biased downward econometric estimates of economies of scale if left uncorrected. As a fix Sokoloff proposed a particular "rule of thumb" imputation for the entrepreneurial labor input. Using establishment level manufacturing data from the 1850-80 censuses and textual evidence I argue that, contrary to Sokoloff's claim, the census did generally include the labor of entrepreneurs if it was economically relevant to do so, and therefore Sokoloff's imputation is not warranted for these census years. However, I also find that the census did understate the labor input in small relative to large establishments as Sokoloff asserted, but for a very different reason. The census purported to collect data on the average labor input but, in fact, the data most likely measure the typical number of workers present. For very small establishments the reported figures on the typical number of workers are biased downwards relative to a true average but this is not the case for large establishments. As a result, the early censuses of manufacturing did overstate labor productivity in small relative to large establishments but the size of the bias is smaller than alleged by Sokoloff.

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

Introduction

The consensus of opinion among American economic historians seems to be that there must have been economies of scale in nineteenth century manufacturing, if for no other reason than the substantial increase in the average size of manufacturing establishments over the course of the century (Atack 1985). That said, there is little agreement on the precise magnitude of scale economies or the relative importance of the underlying causal mechanisms at play. This lack of agreement stems in part from deficiencies of measurement in the nineteenth century manufacturing censuses which, simply put, did not collect all the information necessary to reach definitive empirical conclusions.[1]

The census takers themselves expressed serious doubts about the data they collected. Francis A. Walker, the superintendent at both the 1870 and 1880 census, for example, described the data reported by the manufacturing censuses on capital invested as "entirely untrustworthy and delusive" and "wholly worthless." Nevertheless, such jaundiced views did not prevent others at the time from making use of the data and drawing inferences from them. Controversy about the meaning and interpretation of the data has continued to the present, sometimes on more subtle issues, but ones which turn out to have important consequences.

Perhaps the best known example of such modern controversy is the so-called "entrepreneurial labor input" problem. Although the issue was first raised (and resolved in a particular way) by Atack (1976, chs. 3, 7; 1977) economic historians were made more broadly aware of the implications of the problem in a classic paper by Sokoloff (1984) which reported

[1] See Fishbein (1973), and Atack and Bateman (1999a) for a comprehensive discussion of the history of the nineteenth century manufacturing censuses, the information collected, and available modern samples from the surviving manuscript schedules.

estimates of Cobb-Douglas (and other) production functions using samples of the 1820 and 1850 manuscript censuses of manufacturing. Like the other nineteenth century manufacturing censuses, those for 1820 and 1850 collected information on the number of individuals working in the establishment, classified by age or gender.[2] According to Sokoloff (1984), the reported count of workers in 1820 and 1850 generally excluded the labor input of the owner if he was a sole proprietor (or owners, if there was more than one); however, based on supplementary information for 1820, Sokoloff believed that hired managers were generally counted. In effect, Sokoloff claimed that the census collected information on the labor input of employees, not the employer(s) (see Atack 1976 for a similar argument).

Because the employer or "entrepreneurial" portion of the labor input was allegedly overlooked, the common census measure of labor productivity, value added per worker, is overstated because the denominator, the labor input, is too small. More importantly, the degree of overstatement is not uniform across establishments with different numbers of workers. To the contrary, the upward bias in labor productivity is systematically greater for establishments with few workers because in such establishments the ratio of the entrepreneurial labor input to that of the other workers is greater than in larger establishments. Failure to correct for this bias, according to Sokoloff, systematically overstates labor productivity in smaller relative to larger establishments, thereby causing the usual measure of scale economies – for example, the sum of the labor and capital coefficients in a value-added Cobb-Douglas production function – to be biased downwards. Sokoloff proposed specific and somewhat different solutions for this problem in 1820 and 1850. Once these were implemented, he was able to establish econometrically the existence of economies of scale. Importantly, this conclusion pertained to

[2] The 1820 census collected information on the average number of boys, young women, and adult males working at the establishment separately. The 1850 census reports separately the average number of male and female workers.

establishments that did not make use of inanimate power – that is, non-mechanized – implying that division of labor, as hypothesized by Adam Smith, must have contributed to the productivity gains.

The early response to Sokoloff's paper was favorable, as evidenced by Atack (1987) who adopted Sokoloff's proposed solution in his estimation of production functions using establishment level data from the 1850-70 manufacturing censuses.[3] In later work, however, Atack changed his mind, arguing that, in general, the census did include the labor input of the entrepreneur if it was economically relevant to do so.[4] To my knowledge, Sokoloff never accepted Atack's criticisms, nor has there been a satisfactory resolution of the dispute, leaving economic historians in the lurch as to which point of view has more merit.[5] More importantly and unlike some debates over measurement in economic history, this is far from a trivial dispute because, as argued in Sokoloff's original paper and as shown here, Sokoloff's proposed adjustment has a very substantial effect on measured productivity in small establishments relative to larger and therefore on the extent of measured economies of scale. It is thus fundamental to our understanding of the "treatment effect" of changes in establishment size and organizational form on labor productivity during early industrialization.

In this paper I revisit the entrepreneurial labor input problem in a systematic way using data from the Atack-Bateman (1999a) samples from the 1850-1880 censuses of manufacturing along with textual material from the instructions to census enumerators. The conclusion I reach has two parts. First, Atack's revised position that the census did generally include the labor

[3] In this regard, Atack's (1987) analysis of the 1850 data differs from his 1977 analysis (see footnote 2 above).
[4] See Atack and Bateman (1999b), a revised version of which was later published as Atack and Bateman (2008).
[5] The issue has also been considered for the case of 1871 Canadian census of manufactures; in particular, Inwood and Keay (2012) conclude that the entrepreneurial labor input was properly counted and that no Sokoloff-like adjustment is necessary.

input of the entrepreneur when economically relevant has much to recommend it with regard to sole proprietorships; in particular, his claim is consistent with the textual evidence from the enumerator instructions as well as statistical evidence on the distribution of establishment sizes.[6] However, this is not to say that the census enumerators always included sole proprietors in the count of workers when they were supposed to, as I also uncover several examples that suggest a failure to do so, but such failures appear to be relatively uncommon.

Second, I extend Sokoloff's (1984) analysis of the effect of partnerships (versus sole proprietor) by using data on organizational form that Atack subsequently added to the original 1850-70 Atack-Bateman samples. Consistent with Sokoloff's findings, I show in a regression analysis that partnerships had higher output than sole proprietorships after controlling for the reported number of employees. However, the effects are not statistically significant (and are relatively small), especially after controlling for industry, location, and capital invested. I conclude that Sokoloff's recommended blanket correction for any possible undercount in 1850 and, by inference, 1860-80 censuses is not defensible.

That said, while it may be inappropriate to apply a correction for the allegedly "missing" entrepreneurial labor input in the particular manner that Sokoloff recommended, he was correct that the labor input was under-enumerated in small manufacturing plants relative to large establishments -- but for an entirely different reason. The relative understatement occurs because the census data on the labor input refer not to the total quantity of labor used in the establishment over the course of the census year or even a literal average of the number of workers, as is often assumed (see, for example, Laurie and Schmitz 1981, p. 73). Rather, my

[6] In this regard, my argument is similar to that sketched in Atack and Bateman (1999b), who also argue on the basis of the distribution of establishment sizes that sole proprietors were generally counted correctly. Atack and Bateman's focus in their paper is on profitability in manufacturing; in particular, they argue that a Sokoloff adjustment leads to estimates of returns for small establishments that are systematically too low.

analysis leads me to conclude that the data generally refer to the typical number of individuals at work at the establishment during the census year, where "typical" refers to the number usually present on a normal day of operation. This typical number does not take into account occasional periods of time when more labor might be at work as well as periods when fewer than typical were present. The failure to incorporate above or below typical numbers of workers would not necessarily introduce bias but, as it happens, there is a bias that is asymmetric with respect to establishment size. In effect, the census data on the labor input are more accurate for larger establishments and less accurate from smaller establishments, with the error being one-sided (too low) for smaller relative to larger establishments. This is the same type of bias that Sokoloff identified, albeit for a very different reason.

In support of this claim, I make use of novel data from the 1880 census that were included in the manuscripts but never tabulated and published. In 1880, as in 1870, the Census ascertained the average numbers of adult males, adult females, and child workers separately. However, the Census also included a new question on the maximum number of workers at the establishment at any point in time during the census year. I create a sub-sample of establishments in 1880 for which, if the census data on average numbers were literally correct, no establishment should have reported a maximum number of workers that exceed the sum of adult male, adult female, and child workers. However, as I show, almost half of the establishments in the particular sub-sample did so report. I argue that the most likely explanation is that these establishments gave the typical number at work in response to the enumerator, rather than a true average – a type of answer that the census instructions to enumerators permitted. I go on to show that the likelihood of this happening was decreasing in establishment size, which is the same pattern of bias alleged by Sokoloff.

The 1880 data also allow me to estimate whether the difference between the maximum and typical number of workers had a material effect on labor productivity. I show that it did. Under a set of (reasonable) assumptions it is possible to back out an adjustment factor to correct the reported average number of workers. This adjustment factor reduces labor productivity in small establishments relative to large establishments and thereby raises the estimate of the returns to scale, although to a much smaller extent than Sokoloff's proposed adjustment.

The Emergence of Large Scale Manufacturing in Nineteenth Century America

In the early nineteenth century the overwhelming share of the labor force – approximately 76 percent in 1800 – was engaged in agricultural production. Over the course of the century, a shift of labor out of agriculture occurred such that, by 1900, the share in agriculture had fallen by slightly more than half, to 36 percent (Weiss 1986, 1993). Although much of the re-allocated labor went to the service sector, a significant portion went to manufacturing where labor productivity was substantially higher than on the farm, thus contributing to economic growth and rising material standards of living. Moreover, this growth of manufacturing employment in the nineteenth century was not neutral with respect to establishment size. In particular, the average number of workers per establishment increased (Atack 1987). Smaller establishments, especially sole proprietorships, decreased their share of the total number of establishments and of total employment in manufacturing – a process referred to by labor historians as the "displacement" of the artisan shop by the factory.

Table 1 documents the evolution of this process over the period from 1850 to 1880, using the Atack-Bateman (1999) manufacturing samples. Panel A shows the distribution of establishments by the reported number of workers, while Panel B shows the weighted distribution of the same, in which the weight is the number of workers. In effect, Panel A asks about the typical establishment and variation thereof, whereas Panel B asks in what size establishment we would find the typical worker (and variation thereof). In computing the distributions in Table 1 I have taken the sample data as given, except that I have excluded observations based on the same criteria used in previous work with the Atack-Bateman samples.[7]

Panel A of Table 1 shows the mean and median establishment size, and proportion of establishments in four size categories: 1-5 workers, 6-15 workers, 16-100 workers, and more than 100 workers. In the discussion below, I will refer to establishments with sixteen or more workers as "factories". While any specific cut-off, of course, is arbitrary, the substantive patterns evident in the table do not change for reasonable variations in the cutoff.

The basic finding of Panel A is that the distribution of establishments shifted over time towards larger firm sizes – that is, the share of establishments in the smaller two categories (1-5 workers and 6-15 workers) was declining over time. As a result, the mean number of employees

[7]See, for example, Atack, Bateman, and Margo (2008). In particular, to be included in Table 1, an establishment must be in the (1) "national" Atack-Bateman sample (which is self-weighting for 1850-70) (2) have positive values of employment, capital, value of outputs and inputs, and value added. I have also excluded establishments whose calculated rate of return on capital investment seems too high or low to be credible, as well as establishments for which industry is not reported as well as a few extremely large establishments that, in my opinion, have employment totals that are in error. In addition, I require that gross value of outputs exceed $500 in real 1850 dollars in each census; the corresponding cutoffs are $500 in 1850, $518 in 1850, $848 in 1870, and $576.50 in 1880. These cutoffs are derived using an aggregate price index for manufactured goods developed by Atack, Bateman, and Margo (2004). Also, as discussed, for example, in Atack, Bateman, and Margo (2004) it is necessary to re-weight the 1880 data in an attempt to correct for the under-representation of establishments in so-called "special agent" industries. These establishments were larger than average. The reweighting procedure is imperfect and does not fully correct for the under-representation of larger establishments; hence, the average and median establishment sizes in 1880, and the proportion of establishments with 16 or greater employees, are biased downwards, even after reweighting.

per establishment increased by 46 percent between 1850 and 1880 (from 8.4 to 12.3 workers). Growth in mean establishment size was especially rapid in the 1860s, with some decline in the 1870s as the steady decline in the smallest establishments (1-5 workers) was arrested and the share in the next size class (6-15 workers) actually increased.[8]

Among establishments employing 16 or more workers, growth was especially rapid at the high end – the fraction of establishments employing 100 or more workers nearly doubled between 1850 and 1880. That said, as late as 1880, a very healthy majority of establishments (69 percent) employed between one and five workers. Because the smallest establishments continued to dominate the numbers, the median establishment size of 3 workers did not actually change between 1850 and 1880. Clearly, in terms of the frequency that such establishments appeared in the census rolls, the artisan shop was still dominant numerically in 1880, as it had been in 1850.

Panel A views the size distribution from the vantage point of the establishment as the unit of observation – one establishment, one observation. But the picture is potentially very different from the vantage point of workers. Suppose, for example, that there are two establishments, the first employing a single worker, and the second employing nine workers. The average number of workers per establishment is five but the median sized establishment, when observations (establishments) are weighted by the number of employees, is nine. That is, more than half of the workers in this example are employed in the larger establishment with nine workers. This distinction may be important historically, because it is not known precisely when a majority of American manufacturing workers first began to be employed in factories, as I have defined them,

[8] The estimates of the mean size of establishments are slightly higher than those implied by the published census volumes. For example, the mean size in 1850 according to the published census was 7.8 workers compared with the estimate of 8.4 shown in Table 1. However, the time pattern in the estimates is the same as in the published volumes, in particular, in showing a substantial increase between 1860 and 1870, and a decline in the 1870s.

because the published census data as tabulated provide no information on the underlying distributions.

Panel B shows the size distribution weighting each establishment by its reported employment.[9] The panel demonstrates that, as early as 1850, a majority of workers, 62.6 percent, were employed in establishments that are deemed to be "factories". By 1880 the factory share had risen to 72.7 percent. As the columns of Panel B make clear, growth in the factory share was not uniform across the distribution of 16 workers or more but rather was concentrated in the largest establishments of 100 workers or more, whose proportion of employment grew from 29.5 percent in 1850 to 40.5 percent in 1880. Thus, while the median establishment (Panel A) remained fixed at 3 workers from 1850 to 1880, as a result of the shifts just described, the median worker's establishment size (panel B) rose from 37 workers in 1850 to 60 workers in 1880, a 62 percent increase.

The increase in size evident in Table 1 could be due to shifts in industry composition or location. To determine if this was the case, I estimated a panel regression in which the dependent variable takes the value one if the establishment was a factory (16 or more workers), and zero otherwise. I included dummy variables for the census year, three-digit industry (SIC) code, the state in which the establishment was located and urban status. The left-out census year dummy was 1850 and observations are weighted by reported employment. The coefficient of the 1880 census year dummy in this regression was 0.095, or 9.5 percentage points, very similar to the increase over the same period shown in Panel B of Table 1 (10.1 percentage points).

[9]I do not report the weighted means in Panel B as these are highly sensitive to the extreme values of the distribution of employment in any given year. The medians, however, are not sensitive in this sense and, as discussed in the text, these show a marked increase in establishment size viewed from the vantage point of workers.

Thus, while some compositional shifts played a role, the great majority of the shift of employment towards larger establishments was a general phenomenon.[10]

For many, perhaps most economic historians, the shift of labor towards larger scale production is *prima facie* evidence that economies of scale were present (Atack 1985). Fundamentally, economies of scale arise through division of labor and/or the use of indivisible inputs, which could be physical capital or access to finance. There is no direct evidence of division of labor in the nineteenth century manufacturing censuses, only indirect evidence suggested by differences in the demographic composition or in average wages by establishment size, and by the positive association between the incidence of factory production and transportation improvements (or as Adam Smith famously remarked, "the division of labor is limited by the extent of the market"); see Goldin and Sokoloff (1982); Atack, Bateman, and Margo (2004); Atack, Haines and Margo (2011); and Katz and Margo (2013). Evidence of indivisible physical inputs is suggested by the fact that the diffusion of the steam engine was positively correlated with establishment size, and that larger establishments generated greater labor productivity gains by using steam than did smaller establishments; see Atack, Bateman, and Margo (2008).

It is one thing to observe that scale economies were probably present because establishments were getting larger. It is another thing entirely to measure their size. The extent of scale economies can be estimated parametrically by specifying a cost function or a production function. Because the census data provided at best limited information on costs, scholars have generally tried to estimate the degree of scale economies from the production function. The

[10]I also estimated a similar regression in which the dependent variable took the value one if the establishment employed more than 100 workers, zero otherwise. The coefficient of the 1880 dummy in this regression is 0.127, or 12.7 percentage points, compared with 11 percentage points growth shown in this share in Panel B of Table 1 between 1850 and 1880 (40.5 percent – 29.5 percent = 11 percentage points)..

first such studies were by Atack (1976, 1977) who reported estimates of production functions in 1850-70 by industry-region cells. Using a variable scale parameter specification popular in econometrics at the time, Atack (1977) concluded that there were pervasive economies of scale present in 1850 but these were exhausted at relatively low levels of output and, consequently, in just 5 of the 14 industry regions cells was the typical establishment operating in the range of increasing returns. By 1860, however, the corresponding figure was 9 of 14 cells, suggesting that the optimal plant size was increasing before the Civil War.

Laurie and Schmitz (1981, pp. 74-75) estimate Cobb-Douglas production functions primarily using manuscript census data for Philadelphia in 1850 and 1880 and also some supplemental data for textiles in 1870. Laurie and Schmitz assume that workers reporting zero employees had, in fact, one employee, but this was the only adjustment they made to reported labor input. In 11 of 17 industries in 1850 and 13 of 17 industries in 1880 Laurie and Schmitz reject the hypothesis of increasing or even constant returns to scale. Based on their regressions Laurie and Schmitz argue that American manufacturing in the nineteenth century was not characterized by scale economies, but rather the opposite. Later in the paper I reproduce Laurie and Schmitz's general finding using the Atack-Bateman national samples, and also show that it is reversed if Sokoloff's adjustment is applied to the data.

Sokoloff (1984) is next in the literature but I defer detailed discussion until the next section except to note that Sokoloff argued that there were economies of scale in non-mechanized production in both 1820 and 1850, but these tended to be exhausted at relatively low levels of output, on average, a conclusion similar in substance to Atack's (1977). The implication of this finding, as previously noted, is that division of labor must have played some role in generating labor productivity growth in manufacturing after 1820, but really substantial

scale economies had to await fundamental advances in power transmission and associated machinery.

Finally, Atack (1987) is a comprehensive attempt to assess the extent of economies of scale using the census samples for 1820-1870 as they existed at the time and subsequently extended by Atack and Bateman (1999a). As previously noted, Atack (1987) embraced Sokoloff's adjustment for the entrepreneurial input, although he later had second thoughts. For the majority of industries in every census year that he examined, Atack found efficiency advantages to large scale production – economies of scale – relative to small-scale production – artisan shops. Atack accounted for the persistence of small establishments by noting that many served markets that were protected from competition from more distant competitors by high shipping costs. Improvements in internal transportation and the diffusion of new technologies, such as steam, caused their market share to erode over time.

The studies I have reviewed thus far have relied on econometric estimates of production functions to assess the extent of economies of scale in nineteenth century American manufacturing. An alternative approach is the so-called "survivor method" in which the central concept is the "minimum efficient scale of production (MES)" – the smallest size establishment such that establishments larger than this were increasing (or non-decreasing) their share of aggregate production over time. If the MES is increasing over time, the presumption is that economies of scale are present. This approach was applied to nineteenth century US manufacturing data by Atack (1985). Atack frames his paper in terms of the debate over the "origins of the modern corporation" as told by Chandler (1977). According to Chandler, truly large scale production emerged late in the nineteenth century in response to fundamental changes in technology that were not in place until well after the Civil War. However, Atack (1985, p.

47) shows the typical plant at the end of the century was "little different from the scale required of an efficient plant in 1870" and that, with the exception of a few industries, the long run growth in establishment size can be interpreted as a historical drift towards an equilibrium structure whose fundamental causes were put in play much earlier in the century.

Outside of the United States, the measurement of historical economies of scale has received the most attention by far in the French case. France is interesting because of the famous hypothesis by Landes (1949, 1954) that nineteenth century French manufacturing establishments were "too small" relative to optimal and that by failing to capture unexploited economies of scale, French economic growth suffered compared with, say, England. Economic historians have evaluated the Landes hypothesis by attempting to estimate the degree of economies of scale using the two relevant French censuses from the nineteenth century, 1839-45 and 1861-65. Using a cost function approach applied to the census of 1861-65 Nye (1987) found little evidence that there were economies of scale left to be exploited. However, Siscic (1994), using a production function approach, does find evidence of unexploited scale economies in sectors where the average establishment size was small. The most recent study, that of Doraszelski (2004), is somewhere in the middle, arguing that there were some unexploited scale economies early in the nineteenth century but not in the second half of the century.

The Entrepreneurial Labor Input Problem

The measurement of economies of scale requires accurate information on factor inputs and outputs. In particular, if any inputs are systematically under-reported in small versus large

establishments, an econometric analysis uncorrected for this might show evidence of decreasing returns even if the true production function exhibited increasing returns.

Sokoloff (1984) argued that just such a problem afflicted two nineteenth century manufacturing censuses that he was analyzing, 1820 and 1850.. Sokoloff was interested in measuring economies of scale in both years and, specifically, to determine if such economies were present in establishments that were "non-mechanized" or whether inanimate power was a pre-requisite to obtain productivity gains. If economies of scale were present in non-mechanized establishments, pure division of labor must have been the reason whereas, in mechanized establishments, (quasi)-fixed costs or other factors that made it easier for larger establishments to acquire powered machinery might be the source of the apparent economies, even with division of labor.

Sokoloff (1984) argued that the labor input of entrepreneurs was not properly measured by the census in 1820 or 1850. For 1820, there is sufficient information reported to distinguish establishments that were sole proprietorships versus establishments that were two-person partnerships. Sokoloff computed the difference in value added between sole proprietorships and two-person partnerships, controlling for the reported number of workers (but nothing else), for establishments reporting up to six workers. At each level of reported employment, output per worker was higher in partnership firms than in sole proprietorships. From this difference he inferred that the average number of workers reported in the census did not reflect the labor input of both owners in the partnership, relative to the sole proprietorship, and that it was necessary to inflate the labor input in small establishments because otherwise, as he observed, "firms with one worker would have the highest [measured] value added per worker" and this was unlikely to reflect the historical reality (Sokoloff 1984, p. 369, ftn. 16).

For 1820, Sokoloff added one to the count of workers if there was one owner listed or if the name of the firm was "Jones and Company"; two, if two owners listed; and three, if three or more owners were listed.[11] If the establishment was incorporated or a joint-stock company, Sokoloff assumed that it had a manager, and the manager was properly counted. In short, Sokoloff presumed that the owner(s) of the establishment, as a general rule, contributed labor to production but were systematically excluded from the count of workers in 1820. At the time Sokoloff wrote his paper he had no information on organizational form for the version of the 1850 sample that he analyzed, and thus he could provide no evidence similar to that for 1820 to convince the reader that the entrepreneurial labor input was not counted. Instead, he simply assumed that the same type of undercount occurred in 1850, and to correct for it he added one to the count of workers (Sokoloff 1984, p. 375, ftn. 21).[12]

As noted previously, Sokoloff was not the first economic historian to call attention to this alleged problem. In particular, Atack (1976; 1977, p. 344) asserted that "proprietor and salaried managerial personnel" were "almost certainly excluded" in the 1850 and 1860 manufacturing census. Atack's proposed adjustment for this deficiency imputes managerial and supervisory workers based on the "ratio of [such] personnel to all employees reported in the 1890 census … subject to each firm being assigned at least one manager/proprietor[.]" According to Atack, however, the "returns to scale parameter" was "insensitive to changes in the labor input" induced by his adjustment (Atack 1976). . Here, it is important to note that Atack's adjustment is similar

[11] In the 1850-70 samples for which similar information has been inferred from the name of the establishment, some establishments were "family"-owned enterprises. Presumably there were also such establishments in 1820 but, if so, Sokoloff does not describe how he adjusted the labor input for these.

[12] It is worth noting that both Siscic (2004, p. 467) and Doraszelski (2004, p. 265) followed Sokoloff by adding one to their respective measures of the labor input; as best as can be determined from the published article, however, Nye made no adjustment for entrepreneurial labor. It is possible that this may explain why Nye found little or no evidence of scale economies in nineteenth century France while Siscic found the opposite; however, Doraskelski (p. 265) claims that the Sokoloff adjustment does not affect his substantive conclusions regarding scale economies and that a much more important issue is whether output is measured in value-added or gross value terms.

to Sokoloff's recommended adjustment for 1850 (adding one to the count of workers) in that Atack's also adds at least one to each establishment count of workers, but for establishments with a sufficiently large number of employees, the adjustment will be greater than one. Compared with Sokoloff's adjustment, therefore, which adds a uniform number – one – to each worker count, Atack's adjustment is closer to a proportional adjustment, which may explain why he found no substantive effect on the estimated returns to scale – unlike the case with Sokoloff's adjustment, as shown below.

Every economic historian of the United States knows that census data, particularly those for the nineteenth century, are fraught with error. Individuals traced from one population census to next suddenly age by more or less than ten years; magically become literate at an advanced age or else forget how to read and write; declare themselves to be Irish in one census and born in the United States in the next; or change gender before such an operation was medically feasible. It should not surprise us, therefore, that the various economic censuses have their own special ailments.

With regard to the manufacturing censuses, the list of potential pitfalls is long and serious. Information on months of full-time operation is not reported prior to the Civil War, and information on daily hours of operation is not reported until 1880. As such, only in 1880 is it possible to construct even a rough estimate of annual labor input measured in hours (Atack and Bateman 1992; Atack, Bateman, and Margo 2002; Atack, Bateman, and Margo 2003). With the exception of some information on water and steam power, physical measures of the capital stock are not reported; instead, the capital figures, which are reported in dollars, refer to capital "invested" in the establishment, which could be book or market value, or some combination of both (Gallman 1986). Although the evidence is not completely conclusive, most economic

historians believe that the capital figures do not include working capital (Atack 1977; Atack and Bateman 2008). Information on the characteristics of workers is limited to gender in 1850 and 1860; in 1870 and 1880, gender for adults, and children of both sexes. This means, in particular, that the skills and occupations of workers within establishments are not known; as a result, there is only at best, indirect evidence that large establishments were engaged in division of labor.

Given this laundry list of woes – which, to be clear, is just the highlights, not a complete list -- why privilege the entrepreneurial labor input problem? Table 2, which reports parameters of Cobb-Douglas value-added production functions with and without Sokoloff's adjustment for the entrepreneurial labor input, provides one answer to this question. In Panel A of Table 2 I report the Cobb-Douglas scale parameter, which is the sum of the coefficients of labor and capital minus one, along with its associated t-statistic. In column 2 there are no other variables in the regression specification whereas in column 4 I include dummy variables for urban status, state, and 3-digit SIC industry code. As can be seen there is hardly any evidence of economies of scale in Panel A, except to a trivial and statistically insignificant degree in 1860 if no geographic or regional controls are included. Indeed, the overwhelming evidence is for the reverse – decreasing returns to scale. Panel A, in effect, replicates Laurie and Schmitz (1981) for the whole country rather than just Philadelphia and arrives at essentially the same conclusion.[13]

The situation is very different in Panel B where I implement Sokoloff's adjustment, adding one to the count of workers. Now the evidence is strongly in favor of economies of scale.

[13] Atack's (1977) variable scale production function permits the econometrician to estimate the share of establishment operating in the region of decreasing returns. The most substantial evidence of decreasing returns is found for cotton textiles in 1850 in which 28 percent of establishments in the North and 19 percent in the South are deemed to be subject to (local) decreasing returns; see Atack (1977, p. 348).

Moreover, these scale economies are quite substantial, except in 1880 when the additional controls are entered in the regression.

In computing the labor input for their regressions, Laurie and Schmitz (1981) simply added up the count of male and female workers in 1850; and adult male, female, and child workers in 1880. I have followed the same procedure in Table 2. However, most economic historians who have estimated production functions using the census establishment data have preferred to compute a modified total of the number of workers, which weights female and child workers less than male workers – that is, in "adult-male equivalent" units. The presumption is that female and child workers were less skilled, on average, than male workers, and the weight in question is usually computed from, or at least informed by, the wages of female and child workers relative to males. Such an adjustment to the labor input will also have implications for estimates of economies of scale because, as first demonstrated by Goldin and Sokoloff (1982; see also Katz and Margo 2012), the female/child share of workers was increasing in establishment size – larger establishments were more likely to employ women and children than small establishments.[14]

Accordingly, in Panel C, I replicate the analysis in Panel B adjusting the labor input to take account of the use of female and child labor. I also include dummies for urban status, state, and SIC industry code in Panel C. As can be seen, controlling for the use of female and child labor increases the estimated scale parameter even without any correction for the entrepreneurial labor input, as the discussion immediately above suggests it would. However, the increase in the scale parameter is quite modest because the great majority of establishments did not hire women

[14] Behind this adjustment is an assumption that larger establishments engaged in division of labor, substituting less skilled workers and machines for the skilled labor of the artisan. While division of labor in nineteenth century manufacturing per se is not at question, it must be stressed that the manufacturing censuses provide no direct evidence on division of labor.

or children and thus their measured labor input is unaffected by a Goldin-Sokoloff demographic adjustment. Instead, as Panel B shows, even if such a demographic adjustment is implemented it is still necessary to impute entrepreneurial labor *a la* Sokoloff to generate economically meaningful estimates of economies of scale. Central to the original Sokoloff article was the idea that economies of scale could be had without mechanization, through the pure division of labor. By definition, this pure division of labor involves sub-dividing production tasks between workers to increase the average degree of specialization (reduce the number of different tasks performed by each) possibly to the point where each task is performed by a single worker. By contrast, in the artisan shop, there was much less division of labor – the limit being the one-person shop, in which all tasks are performed by a single worker.

By its very nature, therefore, pure division of labor involves an increase in establishment size, as measured by the number of workers. While the existence of division of labor in nineteenth century manufacturing is not a matter for debate, it must be stressed that the extant establishment level data from the nineteenth century manufacturing censuses provide no direct evidence on its prevalence. Hence, as noted earlier, the importance of Sokoloff's paper – if pure division of labor raised labor productivity, as Adam Smith first observed and which countless generations of economics students have been taught – it should be evident in economies of scale.

In Panel D, I replicate the analysis in Panel C but just for establishments that indicated that no inanimate power source (water or steam) was used in production as a proxy for denying any difference in production methods beyond worker specialization. Nevertheless, as shown in Panel D, the crucial effect of the Sokoloff adjustment is clearly evident on the magnitude of the scale parameter. If the adjustment is used, there is strong evidence of economies of scale even for non-powered establishments, except perhaps in 1880.

Is an Adjustment for the Entrepreneurial Labor Input Warranted? Textual and Statistical Evidence

The analysis in the preceding section demonstrates that Sokoloff's proposed adjustment has a quantitatively large and significant effect on the estimated magnitude of the scale parameter in Cobb-Douglas production functions estimated from the census manufacturing samples. However, while it certainly matters quantitatively whether the adjustment is made or not, by itself this is hardly justification to implement the adjustment – the adjustment is justified if evidence, direct or circumstantial, may be found that the census did, in fact, fail to include the labor of entrepreneurs. In addressing this issue, it is useful to divide up the problem into two parts. Part one concerns the enumeration of labor in firms that are identified as sole proprietors while part two concerns the enumeration of the labor input of partnerships and other organizational forms.

One direct way to address whether an adjustment is warranted is to examine the written instructions to enumerators to see if these gave sufficient guidance as to whom to count as a worker (or not). To be sure, economic historians who hang their hat on the census instructions to enumerators do so at their own peril. The instructions were notoriously sketchy and there is no assurance that they were followed in the field – indeed, as I show below, there is evidence in the case at hand that they were not always followed. That said, the textural evidence does suggest that enumerators were expected to count sole proprietors as "workers" – that is, for sole proprietorships, the reported number of workers should be at least "one".

The specific textual evidence comes from the instructions to enumerators that pertain to wages and to which establishments were at risk of being enumerated, and are taken from Wright's (1900) compilation. As to the former, consider for example, the instructions for questions #10 and #11 in 1860, which reads as follows:

> 10, 11. Wages – Under 10, and 11, entitled Wages, is to be inserted the average Monthly amount paid for all the labor of all the hands, male and female, employed in the business or manufacture during the course of the year. In all cases where the employer boards the hands, the usual charge of board is to be added to the wages, so that cost of labor is always to mean the amount paid, whether in money or partly in money and partly in board; and the average number of hands and the average monthly wages to be returned, so that by dividing the latter by the former the result will show the average earnings of individuals. This is also to be included in the individual labor of a producer, working on his own account, whose productions are separately enumerated.

My reading of this paragraph is that, for sole proprietorships ("the individual labor of a producer, working on his own account, whose productions are separately enumerated") meeting the test for inclusion in the census, the labor input should have been reported as exactly one worker.[15] The test for inclusion, which is what the phrase "separately enumerated" at the end of the quotation is referring to, was specified in 1850 and 1860: separate enumeration was to occur as long as "annual productions" exceeded $500.00 in gross value. This was further clarified in 1870:

> The smallest shop must not be omitted, provided the production reaches

[15] Certainly there is no doubt that sole proprietorships were at risk of being enumerated because, at the top of the census form in, say, 1850, the "owner of the establishment, or business inquired into, either individual, company or corporation" was to be inserted.

> $500 annually, including the cost of materials. It is believed that but few shops which employ the labor of one able-bodied artisan, fall short of this limit at the present prices of labor. Assistant Marshals will take pains to reach all the productive establishments, large and small, within their subdivisions. It is not necessary that there should be a distinct shop to constitute an establishment of productive industry in the meaning of the law. A room finished off in the barn, or a chest of tools kept in the corner of the house, may constitute a distinct establishment, provided the artisan does not habitually work in any other shop which could be separately enumerated.

In sum, the census made provision for counting sole proprietors properly as workers if the owner, "working on his own account", was sufficiently productive to have the establishment enumerated as a separate entity.

Data on the distribution of establishment sizes from the 1850-70 Atack-Bateman samples supports the conclusion that the labor input of sole proprietors was counted properly most of the time. As Atack and Bateman (1999b) note, if such workers were <u>not</u> counted even when their establishments met the condition to be enumerated -- $500 in gross value – we should expect to see a great many zeros in the distribution of employment across establishments, because sole proprietorships were ubiquitous in nineteenth century manufacturing and many of these (like today) were single-person operations, especially in blacksmithing and the skilled trades.

There is no question that there are "zeros" in the distribution of workers hired in the Atack-Bateman samples. These are more common after the Civil War – 4.2 percent in 1870 and 5.2 percent in 1880 – possibly because the Census seems to have been more relentless in canvassing the "smallest shops", as the above instruction to enumerators suggest. Before the

War, however, zeros are uncommon -- in the 1850 and 1860 Atack-Bateman samples, less than one percent of the establishments report zero workers.[16]

Obviously, zero is the wrong answer for any firm of any size. For sole proprietorships reporting zero workers, therefore, a plausible strategy is to impute one worker.[17] For sole proprietorships listing one worker, however, it is not credible to change the one to two workers, because then the distribution of establishment sizes will have too few sole proprietors.[18] This is the fundamental problem with Sokoloff's blanket adjustment

That said, one can find instances of sole proprietorships with one worker in which the "worker" was probably not the proprietor. In the 1870 Atack-Bateman sample, there are six establishments that are identified as sole proprietorships with one worker – except that the worker is a child. Almost surely these are errors – the worker was an apprentice, not the owner.[19] For these establishments, the count of workers could be changed to "two" but were this not done, the substantive effect is minimal because there are very few such observations relative to the total sample size.

Thus, on the basis of the textual and distributional evidence, it would appear that the labor input in sole proprietorships was enumerated correctly, at least most of the time.

[16] The exact proportions are 0.65 percent in 1850 and 0.81 percent in 1860 – that is, in both cases less than one percent of establishments report "0" as the number of workers. These figures pertain to the Atack-Bateman "national samples" and thus are nationally representative of the population of surviving manuscript schedules of manufacturing.

[17] See Laurie and Schmitz (1981). Alternatively, one can simply exclude establishments with missing data on any relevant variable, such as employment; this is the procedure followed by Atack, Bateman, and Margo in their various papers.

[18] See Inwood and Keay (2012) for a very similar analysis in the case of the 1871 Canadian census of manufactures. Inwood and Keay also argue that, in the case of the Canadian census, the enumerator instructions called for counting the labor input of entrepreneurs when it was economically relevant. Inwood and Keay find little evidence, as well, of establishments reporting zero workers in the Canadian data.

[19] One can also find sole proprietorships in which the one worker is female. It is less clear whether these are errors.

However, this leaves open the question of whether the census enumerated the labor input correctly in partnership or other types of establishments. .

In particular, what are we to make of Sokoloff's observation that, in the 1820 census, establishments with two owners had higher value added per worker than sole proprietorships, controlling for the reported number of workers? Sokoloff showed this was the case in 1820 but was unable to investigate this issue for 1850. Subsequently, Atack added the information on ownership to the 1850-70 samples, so it is possible to replicate Sokoloff's 1820 analysis for the later census years.

Table 3 reports the coefficients of a dummy variable for dual (two-person) partnerships in a regression of the logarithm of value added for establishments with one reported worker. For each year I show the coefficient and its associated t-statistic with no additional controls in the regression, and with controls for location and industry, and finally, capital invested. The results with no controls are intended to replicate the specification used by Sokoloff, which simply compared output per worker between the two organizational forms, holding the reported number of workers constant.

For 1850-70 I find the same general pattern that Sokoloff did – namely, that output is higher in partnership establishments versus single proprietorships, even though ostensibly these are all establishments with just one worker. That said, all but one of the coefficients is statistically insignificant at conventional levels; and even this is no longer the case once controls for industry, location, and capital invested are included. Note, as well, that the inclusion of these controls generally reduces the magnitude of the partnership effect, which suggests that the partnership dummy variable could easily be measuring factors associated with higher output not

controlled for in the regression but which had nothing to do with a higher level of the labor input. On the basis of Table 3 there is no reason to suppose that the census properly enumerated the labor input in sole proprietorships but failed to do so in partnerships or other types of firms.

The Meaning of the Labor Input in the Nineteenth Century Manufacturing Censuses: New Evidence from 1880

I have argued that the textual, distributional, and statistical evidence does not favor adjustment for entrepreneurial labor as advocated by Sokoloff (1984). That said, there is evidence that the Census understated the labor input in small establishments relative to large – the same pattern alleged by Sokoloff, although for a very different reason.

In 1880 the census asked a question about employment which it had not asked previously and, to the best of my knowledge, has not asked since. This question ascertained the maximum number of workers employed at the establishment at any point in time during the census year. To understand why the answers to this question are useful in shedding light on whether a Sokoloff-like adjustment to the measured labor input is defensible, it is necessary to look again at the instructions to enumerators, this time at those pertaining specifically to the measurement of the labor input.

Consider, for example, the relevant instructions in 1850 and 1860. "These numbers" – male and female workers – "are to be estimated either by an average of the whole year, or by selecting a day when about an average number was employed, and inserting the number on such

day as the average. You will observe that the enumeration of hands is not to apply to any particular day but to express the average number employed throughout the year". The instructions here were not changed for 1870 and 1880, except that in both years separate information was collected on adult males, adult females, and child workers.

Although it is certainly possible that some establishments reported a literal average of the number of workers, the answers to the additional question in 1880 suggest, rather, that the numbers reported were "typical" – that is, workers present on a day "when about an average number was employed". To see this, consider the subset of establishments in the 1880 Atack-Bateman with one reported worker who was an adult male that claimed to have been in operation for the full year (twelve full-time equivalent months). It is possible to identify such establishments because the 1880 census also asked (in detail) about months of full-time equivalent operation. By definition these establishments did not shut down at any point during the year so they could not have experienced a period of time during which the number of workers fell (temporarily) to zero, the only integer below the average reported (one). If the one worker reported was the literal average, therefore, there should not be any establishments that report a maximum number of workers greater than one. Despite this, 43 percent of these sole proprietorships claimed to have had a maximum number of workers more than one employed at the establishment during the year (fully 15 percent reported a maximum of three or more). Conditional on reporting a maximum greater than one, the mean was 2.5 and the median was two.

The reporting of a maximum number of workers exceeding the average was not confined to the subset of one worker full-year establishments just described. As the first row of panel A of Table 4 shows, substantial fractions, just short of a majority, did so, regardless of size. But as

the second row of panel A shows, the percentage difference between the maximum and the average declined sharply with the average number of workers. The potential for bias, in other words, is greater for the smallest establishments.

In Panel B, I report coefficients from a regression of ln (value added) for establishments in 1880 that operated for the full twelve months of the year. As in Panel A I only focus on establishments that hired adult males, in order to avoid complications that arise because of adjustments to worker characteristics. The regressions include dummies for urban status, state and SIC industry code.

In the first column, I include the maximum number of workers in the regression for one worker establishments, but do not control for capital invested. The marginal impact of an additional worker hired in this manner is to add about 8.8 percent to value added, and the effect is statistically significant. If we make the assumption that these additional workers were as productive per day employed as was the one average worker year round, the coefficient can be interpreted as the fraction of the year that the maximum was employed – in this case, 8.8 percent of the work year, or about one month. Economically speaking, this seems quite reasonable – long enough to make a difference to measured output, but not long enough to warrant inclusion as a "typical" worker. .

The current version of the 1880 Atack-Bateman sample does not include information on the organization of the firm, as do the 1850-70 samples. However, if some of the firms reporting a maximum greater than two were partnerships, it stands to reason that they would have more capital invested, which could account for some of the effect evidence in column 1.

As column 2 shows, when I control for capital invested, the effect diminishes to about 6.4 percent, but it is still statistically significant.

The remaining columns in the table repeat the same regression for establishments of larger size. For these regressions I also include the log of the number of average number of workers, since the bin sizes include establishments of different size, unlike the first two columns. As can be seen, having a maximum greater than the average contributes in a statistically significant way for establishments with 2 to 5 average workers, and even 6-15 workers. But the effect is declining in size, and is entirely absent for establishments with 16 or more workers.

The fact that the maximum number hired contributes less to measured output as size increases, eventually vanishing, is consistent with an asymmetric "measurement error" interpretation. For a firm with say a reported average of 20 workers and a maximum of 24, it is entirely possible – indeed probable – that at some point during the year, fewer than 20 persons were at work. That is, even if the larger establishments were also reporting the "typical" number of employees instead of a true average, the typical number for a large establishment is a better estimate – less biased, and possibly not at all – than it is for a small establishment.

Obviously, the evidence in Table 4 pertains solely to 1880; it is simply not known whether the same phenomenon was present in the early years. However, because the wording of the question did not change and because there is some association between organizational form and value added, as shown in Table 4, it is quite plausible that the same bias is present in 1850 through 1870.

In summary, Sokoloff (1984) was correct that some adjustment to the reported labor input in manufacturing is needed, but was wrong about the reason. The adjustment is greater in size

for smaller establishments, and thus serves to lower the level of labor productivity in such establishments relative to larger ones. The size of adjustment, however, is much smaller than Sokoloff advocated and, consequently, any correction for it has a much smaller effect on the estimated returns to scale.

A simple way to make this point is to add the term for the bias in percentage terms, Maximum Labor – Average Labor)/Average Labor, to the conventional Cobb-Douglass value added production for 1880, and to compare the returns to scale parameter with and without this additional variable. For this exercise, I adjust the labor input for its demographic composition, as in Panel C of Table 2, and I also include dummy variables for urban status, state and SIC industry code. If the bias term is not included the returns to scale parameter is -0.089 (as shown in Panel C of Table 2). If it is included the returns to scale parameter is -0.072. Thus correcting for the asymmetric measurement error in the labor input increases the estimated scale parameter by about 19 percent (in absolute value), but the parameter is still significantly less than one, unlike the result of applying the full Sokoloff adjustment (see Panel C of Table 2).

Concluding Remarks

In a famous paper Sokoloff (1984) argued that the labor input of entrepreneurs was not properly counted by the Census and, therefore, labor productivity in small establishments, particularly those not using inanimate power, was considerably overstated. He proposed an elaborate correction for this alleged bias in 1820 but had to make do with a much simpler adjustment – adding one to the count of workers – in 1850. Initially, economic historians were

favorably disposed to his correction, even for later census years (Atack 1987) but as research on the topic continued, doubts set in.

This paper has taken a fresh look at this controversy using data that was not available to Sokoloff or his chief critic (Atack) at the time. After evaluating textual and statistical evidence for the period 1850-80 I conclude that Sokoloff's particular adjustment cannot be justified. That said, Sokoloff was correct that the labor input in very small establishments was biased upwards, but not for the reason he thought. There is an upward bias because most establishments in nineteenth century manufacturing appear to have reported the "typical" number of workers rather than a true average. Using novel data for 1880, I show that this tendency understates the labor input in small establishments relative to large. A Sokoloff-like adjustment applied to micro-data from the 1880 census to correct for this problem generates a larger returns to scale parameter but it is does not resolve the issue noted by Sokoloff, that, taken at face value, small establishments that were enumerated in the nineteenth century manufacturing censuses appear to be quite productive, complicating the econometric problem of measuring the returns to scale.

As noted elsewhere in this paper, the particular issue focused on by Sokoloff is only one of many problems that vex the estimation of production functions for nineteenth century manufacturing. Economic historians can certainly continue to implement piecemeal solutions to these problems but it may be that this is asking more that the data can truly deliver. Novel data sources with direct evidence on scale economies may be more promising avenues to pursue in our attempt to understand the sources of productivity growth in nineteenth century American manufacturing.[20]

[20] An important example is U.S. Department of Labor (1899), an extraordinary study of differences in labor productivity between hand and machine labor conducted by the Department in the late nineteenth century. The

References

Atack, Jeremy. 1976. Estimation of Economies of Scale in Nineteenth Century United States Manufacturing and the Form of the Production Function. Unpublished PhD Dissertation, Department of Economics, University of Illinois.

Atack, Jeremy. 1977. "Returns to Scale in Antebellum Manufacturing," Explorations in Economic History 14: 337-359.

Atack, Jeremy. 1985. "Industrial Structure and the Emergence of the Modern Industrial Corporation," Explorations in Economic History 22: 29-52.

Atack, Jeremy. 1986. "Industrial Structure and the Size Distribution of American Firms in the Nineteenth Century," Journal of Economic History 46: 463-476.

Atack, Jeremy. 1987. "Economies of Scale and Efficiency Gains in the Rise of the Factory in America, 1820-1900," in P. Kilby, Quantity and Quiddity: Essays in Honor of Stanley Lebergott, pp. 286-335. Middletown, CT: Wesleyan University Press.

Atack, Jeremy and Fred Bateman. 1992. "How Long was the Work Day in 1880?" Journal of Economic History 21: 129-160.

Department collected detailed data on each step in the production of a very specific product – manila envelopes, for example – when made entirely by hand versus machine, including the amount of time, the number (and characteristics) of workers employed, and the specific capital goods used. Atack (1987) provides some limited comparisons of hand and machine productivity using these data noting that they imply a much larger effect of scale than the census manufacturing data do. Other than Atack's preliminary analysis, these data have been neglected by economic historians.

Atack, Jeremy and Fred Bateman. 1999a. "Nineteenth Century American Industrial Development Through the Eyes of the Census of Manufactures: A New Resource for Historical Research," Historical Methods 32: 177-188.

Atack, Jeremy and Fred Bateman. 1999b. "Profitability and Firm Size: The Small Firm Effect In the Nineteenth Century," working paper, Department of Economics, Vanderbilt University.

Atack, Jeremy, Fred Bateman, and Robert A. Margo. 2002. Part-Year Operation in Nineteenth Century American Manufacturing: Evidence from the 1870and 1880 Censuses" Journal of Economic History 62: 792-809.

Atack, Jeremy, Fred Bateman, and Robert 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: 170-194.

Atack, Jeremy and Fred Bateman. 2008. "Profitability, Firm Size, and Business Organization in Nineteenth Century U.S. Manufacturing," in J. Rosenbloom, ed. Quantitative Economic History: The Good of Counting. New York: Routledge.

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

Atack, Jeremy; Fred Bateman; and Robert A. Margo. 2008. "Steam Power, Establishment Size, and Labor Productivity in Nineteenth Century American Manufacturing," Explorations in Economic History 45: 185-198.

Atack, Jeremy; Michael Haines; and Robert A. Margo. 2011. "Railroads and the Rise of The Factory: Evidence for the United States, 1850-1870," in P. Rhode, J. Rosenbloom, D. Weiman, eds. Economic Evolution and Revolution in Historical Time, pp. 162-179. Palo Alto, CA: Stanford University Press.

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

Doraszelski, Ulrich. 2004. "Measuring Returns to Scale in Nineteenth Century French Industry," Explorations in Economic History 41: 256-281.

Fishbein, Meyer. 1973. The Censuses of Manufactures, 1810-1890. National Archives and Records Service Information Paper No. 67, General Services Administration.

Gallman, Robert. 1986. "The United States Capital Stock in the Nineteenth Century," in S. Engerman and R. Gallman, eds. Long Term Factors in American Economic Growth, pp. 165-206. Chicago: University of Chicago Press.

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

Inwood, Kris and Ian Keay. 2012. "Diverse Paths to Industrial Development", European Review of Economic History 16: 311-33.

Katz, Lawrence F. and Robert A. Margo. 2013. "Technical Change and the Relative Demand For Skilled Labor: The United States in Historical Perspective." Working Paper No. 18752, National Bureau of Economic Research, Cambridge MA.

Landes, David. 1949. "French Entrepreneurship and Industrial Growth in the Nineteenth Century," Journal of Economic History 9: 45-61.

Landes, David. 1954. "Social Attitudes, Entrepreneurship, and Economic Development: A Comment," Explorations in Entrepreneurial History 6: 245-272.

Laurie, Bruce and Mark Schmitz. 1981. "Manufacturing and Productivity: The Making of an Industrial Base, Philadelphia, 1850-1880," in T. Hershberg, ed. Philadelphia: Work, Space and Group Experience in the 19th Century, pp. 43-92. New York: Oxford University Press.

Nye, John V.C. 1987. "Firm Size and Economic Backwardness: A New Look at the French Industrialization Debate," Journal of Economic History 47: 649-669.

Siscic, Pierre. 1994. "Establishment Size and Economies of Scale in 19th Century France," Explorations in Economic History 31: 453-478.

Sokoloff, Kenneth. 1984. "Was the Transition from the Artisan Shop to the Mechanized Factory Associated with Gains in Efficiency? Evidence from the U.S. Manufacturing Censuses of 1820 and 1850," Explorations in Economic History 21: 351-382.

United States Department of Labor. 1899. Thirteenth Annual Report of the Commissioner of Labor: Hand and Machine Labor, Volumes 1 and 2. Washington DC: USGPO.

Weiss, Thomas. 1986. "Revised Estimates of the United States Workforce, 1800-1860," In S. Engerman and R. Gallman, eds. Long-Term Factors in American Economic Growth, pp. 641-671. Chicago: University of Chicago Press.

Weiss, Thomas. 1993. "Long-Term Changes in U.S. Agricultural Output per Worker, 1800-1900," Economic History Review 46: 324-341.

Wright, Carroll D. 1900. The History and Growth of the United States Census Prepared for the Senate Committee on the Census. Washington DC: USGPO.

Table 1: Establishment Size Distributions in Manufacturing, 1850-1880

Panel A: Establishments as the Unit of Observation

	N	Mean	Median	1-5 workers	6-15 workers	16-100 workers	>100 workers	≥16 workers
1850	5,019	8.4	3	76.2%	15.1%	7.6%	1.1%	8.7%
1860	4,990	9.9	3	74.4	15.9	8.3	1.4	9.7
1870	3,885	13.8	3	68.6	18.1	11.5	1.9	13.4
1880	7,019	12.3	3	69.2	17.9	10.8	2.1	12.9

Note: Source is Atack and Bateman (1999a) national samples of manufacturing establishments from the 1850-80 manuscript censuses of manufacturing. To be included in the table, observations (establishments) must meet certain criteria; see footnote 6. 1880 data are reweighted to correct for under-sampling of establishments in special-agent industries (see footnote 6); sample size in 1880 is the un-weighted number of establishments, weighted sample size is 7,184.

Panel B: Establishments are weighted by the Number of Workers

	N	Med	1-5 workers	6-15 workers	16-100 workers	>100 workers	≥16 workers
1850	42,032	37	21.1%	16.1%	33.1%	29.5%	62.6%
1860	49,590	51	17.4	14.3	31.9	36.4	68.3
1870	49,595	77	12.3	12.0	31.3	44.5	75.8
1880	84,594	60	13.6	13.6	32.4	40.5	72.7

Note: see Panel A for source. Each establishment is weighted by the number of workers. Number of workers in 1850-60 is the sum of male and females; number in 1870 and 1880 is the sum of adult males, adult females, and children. 1880 data are reweighted (see text to Panel A, above).

Table 2: Cobb-Douglas Estimates of Economies of Scale, 1850-1880: Value Added Production Functions

Panel A: Uses Reported Employment and Capital

Sample	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter
Controls for Location and Industry?	No	No	Yes	Yes
1850-1880 Pooled	-0.031	7.04	-0.064	13.44
1850-1860 Pooled	-0.001	0.11	-0.029	3.99
1870-1880 Pooled	-0.054	9.00	-0.077	11.96
1850	-0.021	2.26	-0.054	5.32
1860	0.018	1.97	-0.006	0.58
1870	-0.016	1.45	-0.038	3.09
1880	-0.074	10.41	-0.099	13.12

For yearly sample sizes, see Table 1, Panel A. Pooled: pooled time series cross section regression with year dummies. Dependent variable is logarithm of value added (value of outputs – value of raw materials). Scale parameter: sum of the coefficients of ln (labor) + ln (capital) – 1 from an OLS regression of ln (value added) on ln (labor) and ln (capital). Labor in 1850 and 1860 is the sum of reported male and female employees. Labor in 1870 and 1880 is the sum of adult men, adult females, and children. Value added is value of outputs minus value of raw materials. Controls for Location and Industry: dummy variables for urban status, state, and 3-digit SIC industry code. 1880 data are re-weighted; see text to Panel A, Table 1 and footnote 6.

Panel B: Uses Reported Capital and Reported Employment Plus One (Sokoloff Adjustment)

Sample	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter
Controls for Location and Industry?	No	No	Yes	Yes
1850-1880 Pooled	0.106	19.98	0.066	11.45
1850-1860 Pooled	0.142	18,18	0.114	12.65
1870-1880 Pooled	0.078	10.84	0.047	6.11
1850	0.127	11.31	0.095	7.53
1860	0.155	14.29	0.131	10.27

1870	0.113	8.59	0.086	5.74
1880	0.059	6.93	0.027	2.93

For definitions, see notes to Panel A. Sokoloff Adjustment: adds one to the count of employees to adjust for alleged under-reporting of entrepreneurial labor input. 1880 data are re-weighted; see text to Panel A, above.

Panel C: Uses Reported Capital and Employment Adjusted for Demographic Composition, With and Without Sokoloff Adjustment

Sample	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter
Sokoloff Adjustment?	No	No	Yes	Yes
1850-1880 Pooled	-0.048	9.83	0.098	16.31
1850-1860 Pooled	-0.007	1.16	0.144	15.66
1870-1880 Pooled	-0.064	9.87	0.078	9.75
1850	-0.035	3.27	0.124	9.59
1860	0.020	1.95	0.163	12.46
1870	-0.015	1.16	0.112	7.65
1880	-0.089	11.68	0.057	6.09

For source, see notes to Panel A. Sokoloff Adjustment: adds one to the count of employees to adjust for alleged under-reporting of entrepreneurial labor input. Demographic adjustment: in 1850 and 1860, labor input is men + 0.6*women; in 1870 and 1880, labor input is men+0.5*women+0.3*children. 1880 data are re-weighted, see text to Panel A, above.

Panel D: Panel C Analysis for Establishments Not Using an Inanimate Power Source

Sample	Number of Establishments	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter	Scale Parameter	Absolute Value of T-Statistic of Scale Parameter
Sokoloff Adjustment?		No	No	Yes	Yes
1850-1880 Pooled	13,539	-0.088	14.35	0.079	10.27
1850-1860 Pooled	6,311	-0.037	4.15	0.132	11.83
1870-1880 Pooled	7,208	-0.123	14.46	0.044	4.06
1850	3,268	-0.065	5.20	0.108	6.88
1860	3,063	-0.014	1.05	0.152	9.50

1870	2,486	-0.077	4.74	0.079	7.65
1880	4,722	-0.150	15.12	0.023	1.80

For source, see notes to Panel A. Sokoloff Adjustment: adds one to the count of employees to adjust for alleged under-reporting of entrepreneurial labor input. Demographic adjustment: in 1850 and 1860, labor input is men + 0.6*women; in 1870 and 1880, labor input is men+0.5*women+0.3*children. All regressions include dummy variables for urban status, state, and SIC industry code. 1880 data are re-weighted, see text to Panel A, above.

Table 3: The Difference in Value Added: Sole Proprietorships versus Dual Partnerships, One Reported Adult Male Worker, 1850-70

Year	Sample Size (estab.)	Dual = 1	Controls	Coefficient of Dual Partnership Dummy	Absolute value of t-statistic
1850	992	0.037	None	0.230	2.12
1850			Location + Industry Dummies	0.128	1.08
1850			Ln (Capital) + Location and Industry Dummies	0.130	1.24
1860	985	0.052	None	0.145	1.56
1860			Location + Industry	0.163	1.76
1860			Ln (Capital) + Location + Industry Dummies	0.142	1.59
1870	615	0.031	None	0.319	1.64
1870			Location + Industry Dummies	0.314	1.59
1870			Ln (Capital) + Location + Industry Dummies	0.183	0.99
1870	390	0.041	Ln (Capital) + Location + Industry Dummies, Full Year Establishments	0.160	0.83

Source: see text..

Table 4: Analysis of 1880 Question on Maximum Employment

Panel A: Average versus Maximum Employment, Sample Statistics

Reported Average Number of Workers	1	2-5	6-15	16 or more
Weighted Sample Mean of Maximum Number of Workers	1.646	3.71	11.66	57.2
Weighted Sample Mean of Average Number of Workers	1	2.83	8.96	47.6
Weighted Sample Mean of [(Maximum – Average)/Average] x 100 percent	64.6%	31.7%	29.4%	24.4%
Percent Reporting Maximum > Average	43.1%	38.8%	46.9%	52.0%
Number of Establishments	1,102	1,719	475	207

Source: 1880 manufacturing sample, Atack and Bateman (1999). Standard 1880 sample inclusion criteria apply to this panel; in addition, establishments had to operate for 12 full-time equivalent months and only employ male workers. Data are re-weighted; see text to Panel A, Table 1.

Panel B: Ln (Value Added) Regressions, 1880 Establishments Hiring Only Adult Males, Full Year Operation: By Average Number of Workers Reported

Reported Average Number of Workers	1	1	2-5	6-15	16 or more
Maximum Number of Workers	0.088 (0.019)	0.064 (0.018)	0.038 (0.009)	0.017 (0.004)	0.006 (0.006)
Ln (Capital)		0.169 (0.019)	0.221 (0.016)	0.216 (0.030)	0.247 (0.037)
Ln (# of workers)			0.631 (0.058)	0.618 (0.118)	0.616 (0.103)

Adjusted R-Square	0.262	0.317	0.517	0.452	0.768

Source: see text and Atack and Bateman (1999a). Data are re-weighted, see text to Panel A, Table 1.