U.S. Historical Statistics

Nineteenth-Century U.S. Industrial Development through the Eyes of the Census of Manufactures

A New Resource for Historical Research

JEREMY ATACK
Department of Economics
Vanderbilt University
and
National Bureau of Economic Research

FRED BATEMAN
Department of Economics
University of Georgia

The purpose of this article is threefold. First, it alerts the profession to the availability of new machine-readable census samples on U.S. manufacturing during the second half of the nineteenth century. Second, it addresses some of the questions, both philosophical and practical, that have been raised regarding the use of the manuscript census of manufacturing data in general and our use of them in particular. Third, it takes the opportunity to illustrate why the manuscripts are so much more valuable for research than the traditionally used published census data. These sample databases represent the culmination of several decades' work.

Our manufacturing data come from the manuscripts of the decennial federal censuses for 1850, 1860, 1870, and 1880. These documents, filled out by the census enumerators in the field, were subsequently sent to Washington, D.C., for tabulation and were later returned to the individual states. Not all of them have survived intact: For example, Georgia and Louisiana records for 1850, 1860, and 1870 no longer exist, whereas those for a number of Ohio and Tennessee counties in 1860 and 1870 are missing, including documents for many of the most important counties—Hamilton in Ohio (Cincinnati) and Davidson in Tennessee (Nashville).[1]

Our data are random samples of manufacturing firms drawn from the surviving records of each state and territory and "national" samples based on the random samples. Presented in a common format with uniform integrated coding to facilitate their use, the samples are available through the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan (http://www.icpsr.umich.edu). The samples are also available, for the time being, from http://www.vanderbilt.edu/Econ/faculty/Atack/atackj.htm. Several grants from the National Science Foundation funded collection of these data.[2]

A Brief History of the Federal Census of Manufactures

The U.S. Constitution, art.1, sec. 2, par. 3, mandates a decennial census of population as the basis for apportioning taxation and representation. From the start, however, there was pressure to extend the inquiry beyond the simple constitutional requirement. For example, James Madison urged collecting information about occupation as part of the first population census in 1790. The request was turned down (U.S. Congress. *Annals of Congress* 1834, 1077–78, 1107–8). Similar memorials were subsequently submitted to Congress by Thomas Jefferson, then president of the American

Philosophical Society, and by the Connecticut Academy of Arts and Sciences for the second census in 1800. These, too, were rejected (U.S. Congress 1870, 35–37; Wright 1900, 19–20).

Eventually, however, Congress was forced to take notice of the country's nascent industrial activity after President Thomas Jefferson's Embargo of 1807 threatened to deprive the nation of much-needed manufactures. As a result, the House passed a resolution requesting a report from Secretary of the Treasury Albert Gallatin on how best to foster domestic manufactures. Gallatin protested that he did not have the resources to comply with this request but noted that the third census would afford an ideal opportunity to collect the desired information. Congress agreed and amended the act for the third census on 1 May 1810—just three months before it was to be conducted—requiring that the enumerators render an accounting of the manufacturing establishments and manufactures in their districts, according to specific instructions from the secretary of the treasury. Unfortunately, Gallatin gave them little direction, not even providing enumerators with a standard set of questions or printed forms on which to collect the information (Fishbein 1973, 2). As a result, the responses were not uniform, and there have been serious doubts about the completeness of coverage—concerns that have continued to plague all censuses to the present (see, for example, Bishop 1868, 159). The task of making sense of these imperfect data fell to Tench Coxe—American patriot and one-time assistant to Alexander Hamilton, then serving as Purveyor of Public Supplies in Thomas Jefferson's and James Madison's administrations—whose efforts to present and interpret them have been praised by subsequent generations despite skepticism regarding the underlying data.[3]

In 1815, following the onset of peace in Europe, U.S. products faced heavy competition both in domestic and overseas markets that threatened collapse for much of the nation's manufacturing. In an effort to address this problem, the bill that authorized the fourth census also contained provisions for a census of manufactures to gather information for future public policy decisions. To avoid the earlier mistake, however, investigators developed a detailed questionnaire based on suggestions made by Adam Seybert, a four-term congressman from Philadelphia. Author of *Statistical Annals* (an early attempt at a statistical abstract), Seybert was a member of the leading learned societies whose questionnaire for the 1820 census is unique among the census inquiries in that it embodied open-ended questions as well as the more usual close-ended questions regarding capital invested, employment, inputs, and outputs.[4] The open-ended questions were meant to elicit information and opinion on business conditions and trends and to gauge the impact of foreign competition on domestic manufactures. Coverage of the Northeast, although incomplete, is generally thought to have been quite good, but manufacturing in the South and Midwest was seriously undercounted. Nevertheless, Kenneth Sokoloff (1982) has demonstrated that this census is capable of yielding useful information on early manufacturing by drawing from a sample of some fourteen hundred manufacturing establishments in the Northeast from the census schedules.

Despite continued intense debate over the tariff question in the late 1820s, no provision was made for a census of manufactures in 1830, apparently because of dissatisfaction with the 1820 count. The tariff issue, however, led to the collection of data on manufacturing in 1831–32 by Louis McLane, secretary of the treasury (U.S. Congress 1833). As was the case for responses from the 1820 census, the McLane report responses were not uniform: some questions were open-ended, with incomplete coverage that ignored manufacturing in virtually the entire South and Midwest.

A formal census of manufactures appeared again in the 1840 census, but successive generations of scholars have expressed considerable skepticism about the quality of these data and dismay at their disjointedness. For example, breweries, gunpowder works, armaments manufacturers, printers, and mills reported the quantities manufactured but not their value; manufacturers of machinery, hardware, and many other industries were not asked how much capital was invested in their businesses, and shipbuilders were asked only for the value of ships built. Consequently, as a microeconomic record of manufacturing activities at the time, the 1840 census is seriously flawed. However, before the sixth census is dismissed out of hand, it should be noted that in his 1976 presidential address to the Southern Economic Association the late Robert Gallman (1979, 1009, fn. 8) argued that these data deserve serious consideration once their limitations are recognized.

Beginning with the Census Act of 23 May 1850, formal provision was finally made for a comprehensive enumeration of "all of the products of industry (excepting agricultural, which is provided for in the previous schedule [Schedule 4—Agriculture]) of each producer or establishment" (U.S. Census Office 1853, xxiv). Thereafter, regular censuses of manufactures were taken as part of the decennial population census until 1900, when the census of manufactures was switched to a five-year schedule (beginning in 1904, with every other quinquennial census of manufactures considered part of the regular decennial census). This pattern survived through the 1919 census of manufactures (taken as part of the 1920 census). In 1921, the census of manufactures was switched to a biennial cycle that continued through 1939, when it was suspended for the duration of World War II. Censuses of manufactures resumed in 1947, with other censuses being taken in 1954, 1958, and 1963. Since 1967, a regular sequence of quinquennial censuses of manufactures has been taken in years ending in "2" and "7" (Bohme 1987); since 1949, these censuses have been supplemented by annual surveys (U.S. Bureau of the Census 1967, 735).

The archival collection of manuscript census data is mas-

sive, a fact that traditionally had been not only a strength but also a serious drawback. Sheer size alone contributed to years of scholarly neglect. Earlier investigators who had been tempted to work with these manuscripts were quickly daunted by the huge amount of data facing them. Even the 1820 census of manufactures, a relatively modest affair at a simpler time in our nation's history, fills twenty rolls of microfilm. The situation truly presented an embarrassment of riches to those concerned with U.S. industrialization. Consequently, historical researchers avoided the census altogether or relied on the published data that the Census Bureau had compiled from the manuscripts. This latter alternative, however, is a very poor substitute. The published records lack the rich microeconomic detail present in the original manuscripts; they reveal no specific information on individual companies; they omit certain questions entirely; and they frequently contain compilation errors. Additionally, researchers can only construct analyses requiring distributional data when they use the manuscripts.

Recognizing the potential for misuse of the data and the intrusive nature of the questions, enumerators promised the respondents confidentiality for what they hoped would be truthful, complete answers. For example, the preamble to the instructions to enumerators in 1850 had warned: "The individual facts are confidentially imparted and received, and will only be published, if at all, in connection with and as part of a great body of similar facts, from which it will be impossible to abstract or distinguish those of individual firms or corporations [in the published census aggregates]" (U.S. Census Office 1853, xxiv; 1860, 24 [our emphasis]). These promises, however, merely represented bureau policy rather than law until the authorization of the thirteenth census in 1910. Then, the law explicitly stated that the information on the economic censuses "shall be used only for the statistical purposes for which it is supplied. No publication shall be made by the Census Office whereby the data furnished by any particular establishment can be identified, nor shall the Director of the Census permit anyone other than the sworn employees of the Census Office to examine the individual reports" (U.S. Congress 1911, chap. 2, 1909, sec. 25 9).

For some censuses, the issue was rendered moot by congressional authorizations to destroy the records. On 17 January 1912, Congress was supplied with a "list of useless papers now on file in the Bureau of the Census" (U.S. Congress 1912), including 500,000 schedules from the 1900 Census of Manufactures. A subsequent executive order (16 March 1912) directed that these lists should thereafter also be submitted to the librarian of Congress for comment "as to the wisdom of preserving such of the papers as he may deem to be of historical interest" (Taft 1912), but many records were still destroyed, including the 1919 (U.S. Congress 1930), 1925 (U.S. Congress 1931), and 1927 (U.S. Congress 1934) schedules. Not all the records were lost, however. Among those known to have survived more or less intact are the records from 1929, 1931, 1933, and 1935 (Raff 1998, 58). Moreover, establishment-level data are available in machine-readable form for each of the quinquennial manufacturing censuses since 1963, although the data have been sanitized to prevent identification of the businesses (Nucci 1998, 253).

The quality of the census data can be no better than the information given to the enumerators. Recognizing this simple truth, census officials tried hard to ensure that respondents really understood not only the questions being asked but also how they should be answered. For example, consider the seemingly simple question: "Name of Business or Manufacture." To guide responses to this question, the census provided the following instructions:

> . . . Under heading 2, "Name of Business or Manufacture" is to be inserted the name of each kind of business or of each kind of article produced, concerning which information is to be obtained. The question which the Assistant Marshal should ask is: What is the name of the business you follow, or the name of the principal article or articles you produce? The answer to this question should state the specific name by which the business or article is known in the place where produced. (U.S. Census Office 1860, 25)

For reporting wages in 1850 and 1860, the instructions were especially detailed and precise:

> 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, when 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 are to be returned, so that by dividing the latter by the former the result will show the average earnings of individuals. (U.S. Census Office 1853, xxiv; 1860, 27)

Notice, in particular, the clarity with which at least one goal of this information was stated in the last clause of these instructions.

Even when questions were repeated from census to census, the wording sometimes underwent changes and refinements reflecting the evolutionary nature of the survey instrument and the quest for "truth." That gradual process is particularly apparent in a comparison between the 1850 and 1860 census instructions. These censuses, along with that for 1870, were conducted under the general terms of the Census Act of 23 May 1850. However, for the later censuses, some significant and substantive modifications were made to the original wording of instructions to the enumerators as well as more subtle changes in phrasing.

For example, at the 1850 census, enumerators were simply advised:

> Under the general heading "*Hands Employed*" is to be inserted under *Male* and *Female*, columns 8 and 9, the average number of each sex employed during the year in the

> manufacture or business. These numbers are to be estimated either by an average of the whole year, or selecting a day when about an average number was employed, and inserting the number on such a day as the average. (U.S. Census Office 1853, xxiv)

In 1860, however, these instructions were elaborated upon as follows:

> 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. Should the business you return be of a transient or temporary nature, as some kinds of fisheries, you should state the number of months during which such operations were conducted. (U.S. Census Office 1860, 27)

In 1870, by way of contrast, no guidance whatsoever was given about how these questions on employment should be answered (U.S. Census Office 1870, 20–26). Instead, the inquiry simply read, "Average number of hands employed—males above sixteen, females above fifteen, children and youth" on the schedule headings (Wright 1900, 314).

There is, however, evidence of some learning by census: a single query of the "total amount of wages paid during the year" was substituted for the questions concerning average monthly cost of labor (Wright 1900, 314). This change greatly simplified matters, especially because enumerators were also instructed to report on the months of operation. In 1880, the census went a few steps further, adding a query about the maximum number of hands employed at any one time during the year. It also inquired about the average daily wage for a skilled mechanic and an ordinary laborer as well as about hours of work and months of operation at various capacity levels (Wright 1900, 315). The instructions themselves provided no elaboration, however (U.S. Census Office 1880, 14–16).

Earlier on, learning-by-doing was necessarily somewhat limited until the census was established as a permanent branch of government in 1902 (at which time the Census Office was renamed the Census Bureau). The changes in instructions and wording were not always for the better, either. For example, in 1870, enumerators were simply told that "the kind of business and character of product should be described as specifically as possible, as for example, fishing hooks, hoisting apparatus, skirt supporters, speaking tubes, etc." (U.S. Census Office 1870, 21–22), whereas the earlier instructions were much more helpful.

The range of questions asked of manufacturers between 1850 and 1880 are shown in table 1. The 1850, 1860, and 1870 censuses of manufactures stand out from the rest with respect to the level of detail regarding raw material inputs and final product outputs. These censuses allowed up to six lines of detail on individual raw material inputs and outputs, providing space for information about type, quantity, and the value of each.

In contrast, the 1880 census, which generally sought much greater detail about employment and wages and source of power, reduced the measurement of raw materials and final products to their aggregate values. Indeed, in 1880, no specific instructions were given to the enumerators collecting manufacturing data beyond the questions on the forms and the admonition that the data on specific industries, such as iron and steel or cotton, were to be collected by special agents, and that other certain industries, such as boots and shoes and leather tanning or paper mills, had to be entered on specific forms. Moreover, special agents were also supposed to collect the manufacturing data for all firms in cities with a population over eight thousand in 1870 (U.S. Census Office 1880, 14–16).

Then, as now, the breadth and depth of coverage by the census were always very much an issue, with particular concern for the possible underenumeration of marginal populations whether of individuals or businesses. In 1850, 1860, and 1870, the cut-off for firms was simply set at $500 output per year; in 1870, however, enumerators were explicitly cautioned:

> The smallest shop must not be omitted, provided the production reaches $500 annually, *including the cost of materials*. It is believed that but few shops, which employ the entire labor of one able-bodied artisan, fall short of this limit at the present prices of labor. . . . It is not necessary that these should be a distinct shop to contribute an establishment of produc-

TABLE 1
Census of Manufactures Questionnaires by Topic, 1850–1880

Topic	No. of questions per census year			
	1850	1860	1870	1880
Location	3	3	3	3
Name of firm	1	1	1	1
Industry	1	1	1	1
Capital invested	1	1	1	1
Labor				
Employment	2	2	3	4
Wages	2	2	1	3
Hours of work	0	0	0	2
Months of operation	0	0[a]	1	5
Motive power	1	2	2	8 or 3[b]
Raw materials				
Type	≤ 6[c]	≤ 6[c]	≤ 6[c]	0
Quantity	≤ 6[c]	≤ 6[c]	≤ 6[c]	0
Value	≤ 6[c]	≤ 6[c]	≤ 6[c]	1[d]
Miscellaneous	0	0	0	industry dependent[e]

Source: U.S. Census Office (1853, 1860, 1870, and 1880).
[a]Instructions called for reporting months of operation if plant was idled for long. In our experience, such reporting was rare.
[b]Eight questions if powered by water; three if by steam.
[c]Up to six inputs and outputs could be itemized. The first four (if there were that many) were to be the most important. The remaining lines were to be for fuel and other items not separately enumerated.
[d]Aggregate value only.
[e]In 1880, special schedules with additional questions were prepared for agricultural implements; paper mills; boots and shoes; tanneries; brick and tile making; sawmilling; flour and grist mills; cheese, butter, and milk factories; and slaughtering and meatpacking. Unlike the special agent reports, these schedules are part of the extant records. Other schedules were also prepared for salt mining and quarrying.

> tive industry in the meaning of the law. A room finished off in the barn, or a chest of tools 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.* (U.S. Census Office 1870, 20)

The results may have fallen short of these praiseworthy goals, but no single early economic data source surpasses the nineteenth-century U.S. federal census manuscripts in quality, in consistency, or in comprehensiveness; from mid-century onward, the census enumerations offer a unique historical record detailing the transformation of the United States from an agricultural to an industrial economy.

Samples from the Manuscript Censuses of Manufactures, 1850–1880

Fred Bateman, James D. Foust, and Thomas J. Weiss (Bateman and Weiss 1981, especially 165–84) began collecting the basic 1850–1870 samples during the era of punch cards and mainframe computers, when computer memory was limited and electronic data storage expensive. Appreciation of these constraints helps explain the limitations and problems with the samples, such as the failure to include firm names. The capacity of each punch card was eighty columns of data, and one of the most pressing questions was how best to fit the manuscript census of manufacturing data within this constraint.

Our solution was to generate three punch cards per firm. The first contained information on firm location, industry, capital, employment and wages, power source, and horsepower (table 2). The second contained information about the types, quantities, and values of the inputs; the third had the same data for outputs. In addition, each punch card contained three identifiers: one for the census year, another to the state sample of which the firm was an observation, and the third designating sequence in the series of three cards per firm. We restricted inputs and outputs to no more than four items. When more than four inputs or outputs were recorded by the census enumerators, the three most important (by value) were coded separately; the remaining inputs or outputs were aggregated together as "miscellaneous" inputs or outputs. Unlike the decision not to record firm names, this loss of information is thought to be of little or no real consequence. These choices imposed constraints on our later sampling efforts as we sought to preserve comparability and to integrate later samples with these earlier efforts.

Early on in the sampling and data-entry process, a decision was made to restrict the data on punch cards to numeric characters. Just like the idea to use coded fixed-length records per observation, this decision was conditioned by Bateman and Foust's prior experience in the late 1960s and early 1970s with their linked agricultural and population sample from the 1860 census, which used variable-length records and allowed alphabetical codes (e.g., "W" = white/Caucasian and "F" = female). The variable-length records have proved troublesome to read, and alphabetic data are more difficult to manipulate.[5]

One other decision was later to prove troublesome: A quick survey of the census records revealed a number of firms producing outputs valued in millions of dollars (seven digits) and output units in the millions and tens of millions (eight digits). Consequently, given the premium on space (a maximum of 80 spaces per card), it was decided that two digits—up to 99 different codes (00 being reserved for null)—would more than suffice to code all conceivable inputs and outputs. This expectation proved true, but only for a while. Certainly, it was true for the southern manufacturing firms in 1850 and 1860. By the time we began coding the more complex and diverse industrial activities in the Northeast, however, we ran out of two-digit codes.

Rather than repunch the work already accomplished, we decided to continue using two-character codes, but this time the two characters comprised a letter of the alphabet followed by a single digit 0–9 rather than two digits. Surely, we thought, this would more than suffice. It did for a while, but by the time we began collecting the 1870 data we had again run out of two-digit codes composed solely of numbers or a letter followed by a number. We then switched to codes that could be composed of two letters as well. Nonetheless, this solution created additional problems because the codes were easily confused during data entry: For example, A0 (A zero) could be mistaken for AO (A oh), and A1 (A one) could be mistaken for AI (A eye). Only recently have we decided that most of these problems have finally been corrected.

We collected the original 1850, 1860, and 1870 samples by first transferring the data from the source (either the original bound volumes or microfilms of these large folio sheets) to worksheets with the fixed length fields for each data item. We blocked out and mapped them into the correct columns for an eighty-column punch card. Items needing codes—states, counties, cities, industries, power sources as well as categories of inputs and outputs—were written out in longhand and coded later, often by someone other than the individual who collected the data. The entries from worksheets were then keypunched, generally by yet another person.

This system, dictated in part by the technology of the day and in part by financial considerations (most notably the availability of work-study students to perform the less-skilled tasks, usually keypunching), provided ample opportunity for error from any number or sources. Mistakes were made in reading the semilegible antique orthography; transcription mistakes occurred in entering the data onto the worksheets; coders miscoded; and keypunchers mispunched and had difficulty in reading the transcriber's and coder's handwriting.

Although it was collected concomitantly with the 1850 and 1860 samples, until recently the 1870 sample did not meet our standards of comfort for accuracy of either the sampling or coding of the underlying manuscript data.

TABLE 2
Census of Manufactures Samples: Variables

dBase field name	Units	Census year 1850	1860	1870	1880	Definition
INDUSTRY	3-digit SIC	√	√	√	√	3-digit code. Where more-detailed classification is impossible, a 2-digit code with trailing zero is used (e.g., 070 = not otherwise specified agricultural services), and 999 is used for conglomerate firms spanning two or more distinct 2-digit SIC industry groups
STATE		√	√	√	√	State: Coded and grouped by region. Alphabetical within regions
CNTYCODE		√	√	√	√	County: Coded
CITYCODE		√	√	√	√	City: Coded
CAPITAL	$	√	√	√	√	Capital invested, in real and personal estate, in business
MEN	#	√	√	√	√	Average number of male hands (over age 16 in 1870 and 1880)
WOMEN	#	√	√	√	√	Average number of female hands (over age **15** in 1870 and 1880)
CHILDREN	#			√	√	Average number of children and youth employed (occasionally also reported in error in 1850 and 1860)
LAB_MAX	#				√	Greatest number of hands employed at any one time during year
M_WAGES_MO	$	√	√			Average monthly cost of male labor
F_WAGES_MO	$	√	√			Average monthly cost of female labor
C_WAGES_MO	$	√	√			Average monthly cost of child labor (mistakenly reported in a number of instances in 1850 and 1860)
SKILL_W	$				√	Average day's wages for a skilled mechanic
USKILL_W	$				√	Average day's wages for an ordinary laborer
WAGE_BILL	$			√	√	Total amount paid in wages for year
MONTHS_OP	Months			√		Number of months of operation reducing part time to full time
MAY_NOV	Hours				√	Hours in an ordinary day's labor, May to November
NOV_MAY	Hours				√	Hours in an ordinary day's labor, November to May
FULLTIME	Months				√	Months in operation on full time
THR_QTRS	Months				√	Months in operation on three-fourths time only
TWO_THRD	Months				√	Months in operation on two-thirds time only
HALFTIME	Months	√	√	√	√	Months in operation on half time only (Idle = 12 – (full time + three-quarters time + two-thirds time + half time))
POWER_TYPE		√	√	√		Power source: Coded
HORSEPOWER	hp	√	√	√	√	Horsepower
WATER_HP	hp				√	Horsepower
STEAM_HP	hp				√	Horsepower
INPUT_1		√	√	√		Kind of input: Coded (includes units of measurement)
INPUT_Q_1	#	√	√	√		Physical quantity per units defined by code
INPUT_VAL1	$	√	√	√		Dollar value c.i.f. consumed (excluding inventory)
INPUT_2		√	√	√		Kind of input: Coded (includes units of measurement)
INPUT_Q_2	#	√	√	√		Physical quantity per units defined by code
INPUT_VAL2	$	√	√	√		Dollar value c.i.f. consumed (excluding inventory)
INPUT_3		√	√	√		Kind of input: Coded (includes units of measurement)
INPUT_Q_3	#	√	√	√		Physical quantity per units defined by code
INPUT_VAL3	$	√	√	√		Dollar value c.i.f. consumed (excluding inventory)
INPUT_4		√	√	√		Kind of input: Coded (includes units of measurement)
INPUT_Q_4	#	√	√	√		Physical quantity per units defined by code
INPUT_VAL4	$	√	√	√		Dollar value c.i.f. consumed (excluding inventory)
INPUTS_VAL	$	√	√	√	√	Aggregate values of raw materials including mill supplies and fuel
OUTPUT_1		√	√	√		Kind of product: Coded (includes units of measurement)
OUTPUT_Q_1	#	√	√	√		Physical quantity per units defined by code
OUT_VAL1	$	√	√	√		Dollar value f.o.b. produced—exclusive of cost of transportation
OUTPUT_2		√	√	√		Kind of product: Coded (includes units of measurement)
OUTPUT_Q_2	#	√	√	√		Physical quantity per units defined by code
OUT_VAL2	$	√	√	√		Dollar value f.o.b. produced—exclusive of cost of transportation
OUTPUT_3		√	√	√		Kind of product: Coded (includes units of measurement)
OUTPUT_Q_3	#	√	√	√		Physical quantity per units defined by code
OUT_VAL3	$	√	√	√		Dollar value f.o.b. produced—exclusive of cost of transportation
OUTPUT_4		√	√	√		Kind of product: Coded (includes units of measurement)
OUTPUT_Q_4	#	√	√	√		Physical quantity per units defined by code
OUT_VAL4	$	√	√	√		Dollar value f.o.b. produced—exclusive of cost of transportation
OUTPUT_VAL	$	√	√	√	√	Aggregate values of products, including jobbing and repairing

SIC = Standard Industrial Classification (see U.S. Office of Management and Budget 1987); c.i.f. = cost, insurance, and freight; f.o.b. = free on board.

Consequently, it was not used much. However, as part of our work collecting the 1880 database, we have conducted a major revision for 1870 that involved inspecting specific possible anomalies in the coded data. As a result, we substantially revised and thus improved the sample, which we now consider to be generally trustworthy and a helpful empirical bridge between the pre– and post–Civil War periods.

Beyond our control are some weaknesses, such as underenumeration and missing records, which are inherent in the basic census collection in 1870. Gathered in the war's aftermath, the enumeration suffered collection problems, particularly for the southern states. Despite its shortcomings, however, our work with the economic information from that census indicates its value as a research source.

The 1850, 1860, and 1870 censuses of manufacturing samples were drawn at random from firms in each state. The sampling scheme was based on the number of establishments reported in each state in the summary census statistics. The goal was a sample size of between two hundred and three hundred firms from each state. The size was thought sufficient to provide adequate microlevel diversity. From this limited perspective, the strategy was successful. However, we eventually realized that, given the localization of specific manufacturing activities and state-to-state differences, we could not easily aggregate the state samples into regional or national ones to paint a broader or bigger picture.

In some but not all cases, one could deal with the problem by using appropriate postsampling weights. This solution was, however, far from elegant, and no easy solution existed when one was interested in the underlying distributions rather than a summary statistic such as the average.

The 1880 sample began life under different circumstances from its predecessors. Whereas the earlier ones had begun as mainframe computer projects, the 1880 database originated on the personal computer. The earlier samples involved recording of considerable handwritten data and consequently risked transfer errors. On the other hand, direct data input was used throughout the 1880 sample. The 1880 database also benefited from our considerable "learning-by-doing" with the three previous samples. The process was streamlined and standardized to minimize error and to smooth the work flow. Personal computers made direct data entry feasible by reducing, though not eliminating, the opportunity for error. Database entry programs also provided for rudimentary checking and cross-checking of data for possible transcription errors. Moreover, the very nature of the 1880 census data itself—particularly the elimination of detail on inputs and outputs—reduced the frequency of errors. As a result, data collection went much more smoothly and quickly, greatly reducing the cost, in time and money, of collecting the sample.

From the start, we conceived of the 1880 databases as producing not only state-level random samples that could describe, at a microlevel, an individual state's manufacturing activity, but also a national sample that would not require postsample weighting.

We have subsequently used the experience we gained from sampling the 1880 census to go back and extend some samples for earlier census years, particularly for the leading industrial states, such as Massachusetts, New York, Ohio, and Pennsylvania. We wanted to create larger national samples of manufacturing activity. For example, we have more than quadrupled the size of the New York samples, in effect potentially quadrupling the size of our national samples whose sizes are constrained by the smallest sampling proportion from among the individual states. We have also taken advantage of the discovery or availability of additional censuses of manufacturing for various states—notably, Illinois, Indiana, Michigan, and Missouri—that were missed during sampling in the 1960s and 1970s.

Early in the project, we used our own idiosyncratic industry codes, but we have since converted to three-digit Standard Industrial Classification (SIC) codes, notwithstanding the ongoing changeover to the North American Industrial Classification System (NAICS).[6] The three-digit SIC level is the most detailed level possible with the 1880 data. In 1850, 1860, and 1870, however, it is possible to classify firms into four-digit (or even finer) industry groups, because we have reasonably detailed descriptions of the products that each establishment produced. We have not taken this additional step, but the necessary data are in the samples for each firm. The absence of individual product descriptions in 1880 has also forced us to use the enumerator's description of a firm's activity (e.g., blacksmithing, flour milling, or an iron foundry) as the basis for the SIC classification, whereas in earlier years we sometimes reclassified firms based on what they actually did, rather than what they called themselves.

This situation was particularly true of establishments described as "blacksmith shops," some of which reported producing only agricultural implements or wagons and carriages. Such establishments were reclassified as agricultural implements manufacturers (SIC 352) and "transportation equipment not classified elsewhere" manufacturers (SIC 379) rather than as blacksmiths (SIC 769). On the other hand, blacksmiths who produced such items along with other products (e.g., horseshoes) were still classified as blacksmiths. We discuss this subject later because blacksmithing illustrates one reason why the census manuscripts are to be preferred to the published census aggregates.

Some of the product or raw material specifications in the 1850, 1860, and 1870 censuses were quite detailed. For example, there were lumber mills that, in addition to using *logs,* used *ash, beech, birch, cedar, chestnut, hardwood, hemlock, mahogany, maple, oak, pine, poplar, spruce, walnut, white oak, white wood,* and *willow.* In conjunction with quantities and values, these detailed descriptions hold the promise of deriving local and regional unit price indexes and "market baskets" for finished manufacturers, semimanufacturers, and raw materials for benchmark dates.

Not all data collected at the censuses were tabulated and summarized by the Census Office for the published volumes. The information on motive power before 1870 was never tabulated by the census nor were the data on hours of work and months of operation in 1880. Thus these data would have been essentially lost but for our samples from the census manuscripts.

At the state level, our samples range in size from as few as 52 firms (the entire population of manufacturing firms in Oregon in 1850) to as many as 1,400 firms in New York in 1880 (table 3). Except for the more industrial states, however, the state sample sizes say more about our commitment to our original sample scheme than to any proportionality between our sample sizes and the extent of each state's manufacturing sector.

These state samples have themselves been sampled and used to generate national samples of manufacturing establishments where the probability of a firm's appearing in the sample was equal across all states for which records are known to exist and could be sampled in each year (table 4).

TABLE 3
Census of Manufactures: State Sample Sizes

Sample: Region/state	Census year			
	1850	1860	1870	1880
South				
Alabama	235	248	275	456
Arkansas	104	112	300	269
District of Columbia	138	148	185	331
Florida	95	76	288	183
Georgia	Missing/not located	Missing/not located	Missing/not located	260
Kentucky	305	275	221	280
Louisiana	Missing/not located	Missing/not located	Missing/not located	476
Maryland	229	222	Missing/not located	290
Mississippi	218	241	325	307
North Carolina	223	216	270	272
South Carolina	284	231	245	235
Tennessee	236	129	107	294
Texas	310	262	254	366
Virginia	276	270	238	242
West Virginia	Statehood 1863	Statehood 1863	208	219
Midwest				
Illinois	233	268	558	516
Indiana	253	295	291	401
Iowa	178	209	349	287
Kansas	Territory 1854, statehood 1861	176	265	314
Michigan	196	242	393	351
Minnesota	Territory 1849, statehood 1858	171	179	266
Missouri	197	189	513	305
Nebraska	Territory 1854, statehood 1867	Missing/not located	Missing/not located	296
Ohio	525	476	562	846
Wisconsin	209	201	186	527
Northeast				
Connecticut	207	199	225	260
Delaware	175	183	239	260
Maine	184	166	195	402
Massachusetts	450	346	305	479
New Hampshire	200	211	193	332
New Jersey	242	237	245	287
New York	1,358	1,020	814	1,400
Pennsylvania	1,168	1,079	983	1,064
Vermont	203	219	187	306
West				
California	93	223	191	304
Oregon	52	288	175	209
Washington	Territory 1853, statehood 1889	Missing/not located	Missing/not located	310
Total	8,776	8,828	9,964	14,202

Note: No samples were taken from Rhode Island; records were unavailable at the time of sampling; some records are now available.

TABLE 4
Census of Manufactures:
Representative Regional Sample Sizes

Sample: Region	Census year 1850	1860	1870	1890
South	1,028	1,019	739	1,469
Midwest	1,156	1,414	1,937	2,832
Northeast	3,318	2,936	2,508	3,634
West	48	73	112	238
Total	5,550	5,442	5,296	8,173

Sample proportions range from as few as 22 firms per thousand in 1870 to as many as 46 per thousand in 1850. Although these sampling fractions are small, the number of firms in each sample exceeds 5,000 and ought to be adequate for most uses.

Indeed, sample sizes are such that, in any one of the benchmark years, there are perhaps as many as 37 three-digit industries in which at least 50 producers are present in our national samples (table 5). And in some industries, such as flour milling (SIC 204), lumber milling (SIC 242), boots and shoes (SIC 314), and blacksmithing (SIC 769), sample sizes are in the many hundreds of firms nationwide. These sample sizes ought to provide ample variety for most uses.

The Accuracy of the Census

Despite the care that went into crafting the survey instruments, the census data are only as good as the weakest link in the chain. Before 1880, the census collection was performed by the U.S. assistant marshals who were enjoined to "visit personally the farms, mills, shops, mines, and other places respecting which information is required" and to obtain "all such information from the best and most reliable sources" (U.S. Census Office 1853, xix). The marshals neither received special training for this task nor did they have any special skills. However, the census manuscripts repeatedly reveal evidence of the care and attention to detail paid by so many enumerators by way of marginal notes and commentaries. Consequently, it seems unlikely that the marshals were the weak links in the chain, although many who have worked with the original returns have sometimes wondered whether the ability to write was viewed as a prerequisite for the job.

The 1880 census was the first to be collected by a professional labor force chosen solely to perform this task, reflecting the growing complexity of the census itself together with ambitions for what the census might become. In particular, the 1880 census was originally conceived as a centennial celebration of U.S. independence. It was a far more ambitious undertaking than any previous count, and its final, incomplete report ran to twenty-two volumes that took eight years to prepare. For the 1880 census, the same enumerators appointed to collect the population statistics were directed to gather the statistics on manufactures, except in cities of over eight thousand persons, in which case a special agent was appointed. When the firm was engaged in the manufacture of cottons, woolens and worsteds, iron and steel, silk, beer and liquor, glass, coke, or mining and drilling, the statistics were to be collected by an industry expert.

Despite exhaustive investigation, no one has been able to locate any of the returns taken by the aforementioned experts (Delle Donne 1973, 15). Thus, any firms from these separately canvassed industries as appear in our 1880 samples are enumeration mistakes and were not generally tabulated from the regular returns by the Census Office.

TABLE 5
National Sample of Industries with Fifty or More Firms in at Least One Census Year

SIC code	Industry	No. of firms per census year 1850	1860	1870	1880
72	Agricultural services	10	3	48	57
175	Carpentry	121	52	334	530
201	Meatpacking	30	50	25	254
202	Dairies	10	2	36	113
204	Flour milling	443	547	406	801
205	Bakeries	82	79	74	223
208	Beverages	67	130	101	58
212	Cigars	45	62	135	241
223	Broadwoven woolens	80	69	46	15
228	Yarn	59	28	13	6
231	Men's clothing	162	156	208	268
235	Millinery	63	72	50	57
242	Sawmills	778	908	587	835
243	Millwork	43	36	99	209
244	Wooden containers	147	133	148	127
251	Wood furniture	201	161	125	226
262	Paper mills	31	27	27	60
271	Newspapers	24	61	36	20
273	Book publishing	16	13	8	118
286	Organic chemicals	52	52	8	17
287	Nitrogenous chemicals	52	3	22	19
311	Leather tanning	314	250	119	117
314	Boots and shoes	563	543	523	660
319	Saddlery and harness	172	158	166	283
325	Brick and tile	81	84	70	176
331	Iron forges and steel	58	48	28	19
332	Iron castings	58	42	51	46
335	Nonferrous metals	5	2	134	226
342	Edge tools	62	31	16	51
344	Sheet metal	99	5	5	10
349	Miscellaneous fabricated metal products	8	145	13	20
351	Steam engines	52	21	60	123
352	Agricultural implements	49	113	53	245
371	Carriages	0	0	0	112
379	Wagons and carriages	237	259	251	90
391	Jewelry	23	34	52	63
769	Blacksmithing	484	367	612	862
999	Conglomerate enterprises	161	107	31	0

In return for collecting the 1850, 1860, and 1870 data, the assistant marshal was to receive fifteen cents in payment for every "productive industry, fully taken and returned" without regard for the time and effort required to secure such information (U.S. Census Office 1853, xix). At that time, the daily wage was about a dollar a day or more, so the compensation to the marshals was not especially generous, particularly in less densely settled areas, when considerable travel time might be required between establishments. Thus, in the more rural areas, payment on a per-firm basis may have led to the neglect of isolated manufacturing activities. Consequently, Francis A. Walker, the census superintendent for the 1870 and 1880 censuses, expressed skepticism that the enumerators had either the skill or the incentive to guarantee the accuracy of the returns.

Walker noted several potential reasons for error, including the following:

- Since the collection of manufacturing statistics was merely incidental to the collection of population data, it was [an] obligation to be discharged as expeditiously as possible;
- Assistant marshals lacked the necessary technical expertise to track down and identify all manufacturing enterprises which may have been hidden or which shared premises with other enterprises (but note the warning given enumerators regarding this);
- The inadequate compensation of assistant marshals;
- Dereliction of duty by a small fraction of assistant marshals resulting in a 50 percent or more under-enumeration;
- The temporary absence of proprietor, ignorance or reluctance to provide information;
- The geographic separation of administrative offices from workshops; and
- The incentive among some producers to declare a value of production of less than $500, thus escaping enumeration. (U.S. Census Office 1872, 372–74, 386)

However, the collected data generally appear to be internally consistent—for example, input and output values are somewhat proportionate, and materials and products are appropriate for a particular line of manufacture—with any gross peculiarities noted in marginal comments.

Enumerators were paid only after their sheets were received and checked for obvious defects in Washington, D.C., where the data were tallied to produce the published summaries. It was also in Washington that much tampering was done to the data. Because we are well aware of data manipulation as well as the obvious care taken by most enumerators in the field, we are skeptical of studies that rely solely on published census data.

Three distinct problems existed after the census enumerations reached Washington, D.C. First, among users of census data, the heavy hands of the "Washington gnomes" are well known. By replacing information reported by the original enumerators with other (more appropriate?) numbers, the Washington compilers often greatly altered the published summary of figures. For example, wages in an entire area were sometimes altered to make the monthly wage uniform across firms. This notion might have fit with those of perfectly competitive labor markets, but it hardly holds true today and was even less likely true then. The basis for this seemingly arbitrary change is unknown. Most experienced users are skeptical of the validity of these alterations, believing the original data to be, on balance, more accurate. Certainly, our comparisons of the original data with the published summaries thereof often provide sufficient evidence to impeach work based on those published figures.

Second, the tallies themselves were not correct. In Florida, in 1850, nine firms (four cotton presses employing 111 men and $89,000 in capital to produce output valued at $30,400 as well as five sugar refineries employing 121 men, 65 women, and $217,150 in capital to produce output valued at $23,100) were not included in the tally of that state's manufacturing.[7] The count for Oregon (among many other errors) missed one meatpacker, one planing mill, one gas manufacturer, one shipyard, and one maker of beehives.[8] We have found numerous similar errors whenever it was possible to check. Indeed, if such errors abound in the tabulations of relatively small states with uncomplicated manufacturing sectors, it seems likely that even more flagrant errors will be present in the larger, more complex states. This finding raises a serious complication in that it eliminates the published census as a yardstick against which our samples might be compared.

Third, the published census contains printer errors—some known, others unknown. The most serious error of which we are aware is in data for California, in 1850. The published volume reports two flour mills producing $754,192 of output, whereas the manuscripts show their combined output to be just $77,950.[9] Moreover, errors do not necessarily offset one another. In California, in 1850, the census overstated the output of the state's lumber mills by $70,400 while underreporting the capital of the two flour mills by $15,000.[10]

By our cataloging this litany of census errors, we do not mean to imply that our own samples are without error. On the contrary, we are certain that errors remain despite our best efforts. Some mistakes arise from faulty transcription and the like. Other inaccuracies arise from our efforts to decipher the writing of persons of sometimes questionable penmanship in the field in ink that is now faded and deteriorating. For example, a zero sometimes looks like ( ) and may be interpreted as 11, thus generating an error of a scale of a factor of 10+ in that item. Fortunately, such errors are often easy to spot when one looks at ratios and various different calculations. Other errors—"8"s that look like "3"s, "7"s that look like "4"s, for example—are harder to ascertain and may well have gone undetected, depending on the placement of the questionable digit in the sequence.

Superintendent Walker suggested that one measure of the error of omission might be a comparison between the statistics of industry and of occupations. In 1860, for example, 112,357 persons reported their occupation as blacksmith even though only 15,720 blacksmithing establishments

were enumerated. In 1870, 141,774 persons gave their occupation as blacksmiths, but only 52,982 blacksmithing establishments were reported. Some differences may be accounted for by the fact that blacksmiths' assistants may have described themselves as blacksmiths. Doubtless many smithies, especially those who were peripheral to other activities, such as farming, or those who were offered only part-time seasonal employment, went uncounted. If a full accounting had been made, Walker estimated that gross manufacturing product in 1870 would have been increased by at least $600 million, that is, by about 15 percent (U.S. Census Office 1872, 373–76).

Interpreting the Data

The census went to great lengths to provide instructions and guidance to the enumerators and to update those instructions in the light of experience. Despite this care, however, some items resisted interpretation and still do. The capitalization figures, in particular, have been heavily criticized. Walker, despite his position as superintendent of both the 1870 and 1880 censuses, called the figures "entirely untrustworthy and delusive" and "wholly worthless," noting that "no man in business knows what he is worth—far less can say what portion of his estate is to be treated as capital." Walker continued, "It is a pity . . . that statistical information . . . of high authority and accuracy should be discredited by association with statements so flagrantly false" (U.S. Census Office 1872, 381–82). Walker repeated these criticisms in his remarks prefacing the 1880 census, giving as an example the case of two otherwise identical shoe manufacturers, each employing two hundred workmen. Whereas one owned the building in which his operations were conducted, reporting capital as $75,000 (the value of stock and machinery plus the building), the other rented his facilities and reported only $25,000 in capital, this being the value of stock and machinery only (U.S. Census Office 1883b, xxxix). The records of these cases are indistinguishable from one another in the census data.

The level of detail found in the census manuscript provides new insight into activities, including the whole process of industrialization and perhaps into the de-industrialization of the South that occurred after the Civil War. Consider, for example, the case of just two industries—agricultural implements (SIC 352) and wagons and carriages (SIC 379). Both were produced by specialized firms, some of which became industrial leaders (e.g., John Deere and Studebaker). Each industry was clearly identified at each census. Thus, manufacturers of agricultural implements are credited with producing $6,842,611 worth of output in 1850, $17,597,960 in 1860, and $52,066,875 in 1870. The output of wagon and carriage makers shows similarly rapid growth from $6,827,451 in 1850 to $26,848,905 in 1860 to $65,362,837 in 1870 (U.S. Census Office 1872, 394–408).

These figures, however, hide much more than they reveal, for each of these products was also produced in establishments that were classified by the census as "blacksmith shops." Blacksmiths were jacks-of-all-trades throughout most of the nineteenth century. They produced a wide range of goods that fully deserve to be called "manufactured products." Blacksmiths used more than 40 percent of the national iron production to make a diverse range of consumer and producer goods, ranging from horseshoes to pots and pans to agricultural implements. Their production of plows, fanning mills, hoes, scythes, knives, and wagons and carriages represents an important, if declining, share of the total output of these goods (see U.S. Census Office 1883b, 927). Blacksmiths thus contributed to agricultural progress and development, especially in the more remote parts of the country. Equally important, the experience they gained when they manufactured these products enhanced the development and diffusion of mechanical expertise and technology.

Rather than considering blacksmithing to be an anachronistic craft activity irrelevant to industrialization, we perceive it as one of the most obvious points of contact between the established agricultural sector and that of nascent manufacturing. Small in scale, technologically uncomplicated, and familiar to farmers, blacksmithing offered the agriculturalist a convenient way to move human and physical capital into the newly developing manufacturing sector, though not as quickly as one might have hoped (Bateman, Foust, and Weiss 1974). Blacksmiths also played an obviously important role in providing capital to the farm sector. Ignoring blacksmiths—many of whom evolved into important implements makers (for example, Deere) or were supplanted by them—distorts the economy's long-term adjustment to developmental opportunities. The census reports published after 1890 omitted blacksmithing from manufacturing, so the fact that blacksmithing at mid-century was a form of manufacturing (just as were lumber milling and many other activities) was essentially ignored.

In the South, where few farm implement manufacturers existed, we estimate that, in 1850, blacksmiths produced farm implements such as plows, rakes, and hoes equal to about a quarter of the output reported by specialized southern agricultural implements manufacturers (table 6). Furthermore, we estimate that blacksmiths also produced about

TABLE 6
Blacksmiths as Manufacturers: Production of Agricultural Implements and Wagons and Carriages

	Agricultural implements (% output)			Wagons and carriages (% output)		
Region	1850	1860	1870	1850	1860	1870
South	25	10	129	9	11	14
Midwest	5	3	1	14	6	5
Northeast	3	2	< 0.1	9	6	3

9 percent of southern output of wagons and carriages. In 1860, they produced about 10 percent of the output of the specialist agricultural implements manufacturers and about 11 percent of the wagon and carriage output.

What is even more revealing, however, is what happened after the Civil War. In all regions of the country *except* the South, specialized product manufacture by blacksmiths of agricultural implements, wagons and carriages, edge tools, and the like continued to decline; in the South, however, the output of these products by specialized manufacturers continued to rise. In the South, blacksmiths actually increased their share of the total production of these kinds of products.

Conclusion

Our newly released national and state samples from the censuses of manufacturing for 1850, 1860, 1870, and 1880 open up unprecedented opportunities for microlevel research into the process and pace of industrialization in the United States in the nineteenth century. These database samples provide the proverbial mountain of evidence for research by economists, economic historians, and others concerned with industrial and economic transformation. Now accessible and readily usable by means of research technology, these samples promise to open new doors for historical research.

NOTES

1. Many missing pages of the Ohio records were used to wrap other manuscripts.

2. The original 1850, 1960, and 1870 samples were collected by Fred Bateman and James D. Foust, of the Indiana University, and Thomas Weiss, of the University of Kansas, under Grants 95-2450, 95-2456, SOC 75-18917, and SOC 75-20034 from the National Science Foundation (NSF). The 1880 data were collected by Jeremy Atack (then at the University of Illinois) and Fred Bateman (then at Indiana University) under Grants SES 86-05637 and SES 86-09392. Atack and Bateman subsequently supplemented and modified the samples, using funds provided by the NSF, Vanderbilt University, and the University of Georgia.

3. North (1899, 265), for example, compares Coxe's estimate of manufacturing (U.S. Congress. *American State Papers.* 1832, 666–812) with the Domesday Book.

4. The 1820 Census of Manufactures is available from the National Archives, Record Group 29.

5. See Bateman and Foust, Agricultural and demographic records of 21,118 rural households selected from the 1860 manuscript censuses, available from ICPSR (Study 09117). A rectangular version of an earlier, uncorrected copy of this sample data set (which also loses some information) is also available from ICPSR as Study 07420. See http://www.icpsr.umich.edu/archive1.html.

6. See http://www.naics.com. The observant reader will note that whereas the SIC codes were a federal government standard administered through the Office of Management and Budget, NAICS is a commercial venture.

7. These data are based on a comparison between our 100 percent sample of Florida manufacturing in 1850 and the published count.

8. These data are based on a comparison between our 100 percent sample of Oregon manufacturing in 1860 and the published count.

9. These data are based on a comparison between our 100 percent sample from California in 1850 and the published count.

10. These data are based on a comparison between our 100 percent sample from California in 1850 and the published count.

REFERENCES

Bateman, F., J. Foust, and T. Weiss. 1974. The participation of planters in manufacturing in the antebellum South. *Agricultural History* 48: 277–97.

Bateman, F., and T. Weiss. 1981. *A deplorable scarcity: The failure of industrialization in the slave economy.* Chapel Hill: University of North Carolina Press.

Bishop, L. 1868. *A history of American manufactures from 1608–1860.* 3 vols. Philadelphia: E. Young.

Bohme, F. G. 1987. U.S. economic censuses, 1810 to the present. *Government Information Quarterly* 4: 221–43.

Delle Donne, C. R. 1973. Federal census schedules, 1850–1880: Primary sources for historical research. *National Archives and Records Service Reference Information Paper No. 67,* General Services Administration.

Fishbein, M. H. 1973. The censuses of manufactures, 1810–1890. *National Archives and Records Service Reference Information Paper No. 50,* General Services Administration.

Gallman, R. E. 1979. Slavery and southern economic growth. *Southern Economic Journal* 45:1007–22.

North, S. N. D. 1899. Manufactures in the federal census. In *The federal census: Critical essays by members of the American Economic Association.* New York: Macmillan.

Nucci, A. R. 1998. The center for economic studies program to assemble economic census establishment information. *Business and Economic History* 27(1): 248–56.

Raff, D. M. G. 1998. Representative firm analysis and the character of competition: Glimpses from the Great Depression. *American Economic Review* 88(2): 57–61.

Sokoloff, K. L. 1982. Industrialization and the growth of the manufacturing sector in the Northeast, 1820–1850. Ph.D. diss., Harvard University.

Taft, W. H. 1912. Executive order no. 1499. *Federal Register.* 16 March.

U.S. Bureau of the Census. 1967. *Statistical abstract of the United States.* Washington, D.C.: GPO.

U.S. Census Office. 1853. *The seventh census of the United States, 1850.* Washington, D.C.: Robert Armstrong, Public Printer.

———. 1860. Instructions to assistant marshals. Washington, D.C.: George W. Bowman.

———. 1870. Instructions to assistant marshals. Washington, D.C.: GPO.

———. 1872. *The ninth census of the United States, 1870. Vol 3. Wealth and industry.* Washington, D.C.: GPO.

———. 1880. *Instructions to enumerators.* Washington, D.C.: GPO.

———. 1883a. *Compendium of the tenth census. Part 2.* Washington, D.C.: GPO.

———. 1883b. *Tenth census of the United States, 1880. Vol. 2. Manufacturing.* Washington, D.C.: GPO.

U.S. Congress. 1833. Documents relative to the manufactures in the United States. 22d Cong., 1st sess. H. Doc. 308. Washington, D.C.: Duff Green.

———. 1870. Report on the ninth census. Vol. 1. 41st Cong., 2d sess. H. Rept. 3 (U.S. Serial Set 1436) 18 January.

———. 1911. Statutes at large of the United States of America, March 1909–March 1911. 61st Cong., 1st sess. Washington, D.C.: GPO.

———. 1912. A list of useless papers now on file in the Bureau of the Census, 62d Cong., 2d sess. H. Doc. 460 (U.S. Serial Set 6325) 17 January.

———. 1930. A list of useless papers . . . Department of Commerce. 71st Cong., 2d sess. H. Rept. 1267 (U.S. Serial Set 9192).

———. 1931. A list of useless papers . . . Department of Commerce. 71st Cong., 3d sess. H. Rept. 2611 (U.S. Serial Set 9327).

———. 1934. A list of useless papers . . . Department of Commerce. 73d Cong., 2d sess. H. Rept. 1728 (U.S. Serial Set 9776).

———. *American State Papers.* 1832. *Documents legislative and executive of the Congress of the United States, Vol. 2. Finance.* Washington, D.C.: Gales and Seaton.

———. *Annals of Congress.* 1834. *The debates and proceedings in the Congress of the United States, Vol. 1.* Washington, D.C.: Gales and Seaton.

U.S. Office of Management and Budget. 1987. *Standard industrial classification manual.* Washington, D.C.: GPO.

Wright, C. D. 1900. *The history and growth of the United States census.* Washington, D.C.: GPO.