NBER WORKING PAPER SERIES
ON
HISTORICAL FACTORS IN LONG RUN GROWTH

FIXING THE FACTS: EDITING OF THE 1880 U.S. CENSUS OF OCCUPATIONS WITH IMPLICATIONS FOR LONG-TERM TRENDS AND THE SOCIOLOGY OF OFFICIAL STATISTICS

Susan B. Carter
Richard Sutch

Historical Paper 74

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1995

Versions of this paper were presented to the Social Science History Association, the Stanford-Berkeley Demography Seminar, and the All-University of California Group in Economic History Conference. We thank members of these audiences, especially Matthew Sobek, Carolyn Moehling, and Daniel Scott Smith for constructive comments and suggestions. Though not a member of one of these audiences, Margo Anderson also offered valuable advice and counsel. Some of the data utilized in this paper were made available by the Inter-University Consortium for Political and Social Research. The data for the *Census of Population, 1880 United States: Public Use Sample (1 in 100 Sample) Computer File* were originally collected by Steven Ruggles and Russell R. Menard of the University of Minnesota, Department of History, Social Science Research Laboratory. The data for the *Census of Population, 1900* were originally collected by the Center for Studies in Demography and Ecology, University of Washington. We gratefully acknowledge the work of Roger Ransom, Tom Tang, Charles Wetherell, and Hongcheng Zhao all of the University of California, Riverside for assistance in handling this data set. We thank the National Science Foundation, the National Institute for Aging, the Institute for Business and Economic Research, University of California, and the Academic Senate of the University of California for financial support at various stages of this research. This paper is part of NBER's research program in the Historical Development of the American Economy. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Historical Paper 74
October 1995

FIXING THE FACTS: EDITING OF THE 1880 U.S. CENSUS OF OCCUPATIONS WITH IMPLICATIONS FOR LONG-TERM TRENDS AND THE SOCIOLOGY OF OFFICIAL STATISTICS

ABSTRACT

We argue that the enumerators' occupational returns from the important census of 1880 were heavily edited prior to publication. The effect was to substantially reduce the number of individuals reported with an occupation. For youthful and older males and all women the editing was so substantial as to qualitatively affect the apparent trend in labor force participation for these groups over time. The stylized facts regarding labor market dynamics during the period of American industrialization and the historical stories constructed around them will now need to be reexamined. We contend that the editing was secretly authorized by Francis Amasa Walker, Superintendent of the *Tenth Census of 1880* and one of the most prominent and decorated economists, statisticians, and public servants in America at this time. While other scholars have identified potential sources of bias in census figures, no one has heretofore suggested that the official statistics of the United States were covertly altered to present a picture different from information collected by census enumerators. If we are correct, the sociology of official nineteenth-century American statistics will require rethinking.

Susan B. Carter
Department of Economics
University of California, Riverside
Riverside, CA 92521

Richard Sutch
Institute of Business and Economic Research
University of California, Berkeley
Berkeley, CA 94720
and NBER

Fixing the Facts:

Editing of the 1880 U.S. Census of Occupations
with Implications for
Long-Term Labor Force Trends and
the Sociology of Official Statistics

Occupation statistics published in the federal census of 1880 have had a disproportionate influence in statistical portraits of the level and distribution of employment in nineteenth-century America. They are as complete as any other nineteenth-century census of occupations.[1] They were collected and compiled by Francis Amasa Walker, Superintendent of the *Tenth Census of 1880* and the most prominent and decorated economist, statistician, and public servant in America at his time.[2] Because Walker placed special emphasis on the collection and reporting of accurate occupation data, the 1880 occupation statistics are generally viewed as the most trustworthy, detailed, and comparable of all those collected in the nineteenth century.[3]

We challenge this portrait of reliability and accuracy.

We believe that the enumerators' occupational returns from this crucial census were heavily edited to reduce the number of individuals with occupations prior to their publication. For youthful and older males and all women the editing was so substantial as to qualitatively affect the apparent trend in labor force participation for these groups over time. The stylized facts regarding labor market dynamics during the period of American industrialization and the historical stories constructed around them will now need to be reexamined. We contend that the editing was secretly authorized by Walker himself. While other scholars have identified potential sources of bias in census figures, no one has heretofore suggested that the official statistics of the United States were covertly altered to present a picture different from information collected by census enumerators.[4] If we are correct, the sociology of official nineteenth-century American statistics will require rethinking.[5]

The 1880 Occupation Statistics and Long-Term Labor Force Trends

According to the standard estimates, industrialization had profoundly different effects on the labor force involvement of youths, older workers, and married women. The participation of youths traced a humped-shape, first rising and then falling in the later stages of industrialization; the rate for older men declined monotonically; and paid employment for married women followed a U-shape, falling to nearly zero levels before beginning the well-known steady advance in the twentieth century.

The standard child labor estimates were developed by Edwards who revised published census figures. Edwards' estimates indicate a substantial increase in the labor force participation of youths between 1870 and 1900. Over this period the employment incidence of boys grew by more than a third from 19.3 to 26.1 percent of the population while the involvement of girls increased from 7.0 to 10.2 percent (Edwards 1943, 97, Table XVIII). After 1900 youth employment rates fell dramatically according to Edwards so that thirty years after the turn of the century fewer than five percent of 10 to 15 year olds were gainfully occupied. Edwards' estimates are shown in Table 1.

The published census figures on the employment of youths had a profound influence on public policy makers, social reformers, trade unionists, and social scientists of the day. The appearance of the first census figures on child labor in 1870 sparked an immediate and vigorous response from labor leaders and reformers who branded child labor as one of the central evils of their day (Abbott 1908, 36; U.S. Bureau of Labor 1910, 33-41; Carroll 1923, 81; Abbott 1947, 261). The *increase* in child labor in each of the federal censuses from 1870 to 1900 prompted contemporaries to look for an explanation in the industrialization which was transforming so many other aspects of social and economic life. The analysis of Abbott (1908) both reflected and shaped the thinking at the time. Quoting a variety of sources from early in the nineteenth century, Abbott located the origins of child labor in the Puritan idea of virtue -- "having 'all hands employed'" (Abbott 1908, 37) -- and the contemporary belief that manufacturing was better suited than agriculture to the employment of children.

> The work of manufactures does not demand able-bodied men, it is claimed, but "is now better done by little girls from six to twelve years old" (Abbott 1908, 24).[6]

In Abbott's view, the expansion of manufactures led to an inevitable expansion of child labor. Abbott also posited a self-reinforcing mechanism which led to an acceleration of the use of child labor even more rapidly than the growth of manufacturing. Echoing the views of labor leaders of the period, she suggested that child labor was the result of an "insufficiency of the man's wages." She added, though, that child labor "was certainly as much cause as effect" in depressing adult male wage rates (Abbott 1908, 36).

Historians have largely accepted the conclusions of these contemporaries and the implication of Census figures -- or Edwards' more polished version of them -- that child labor became increasingly

common in the latter part of the nineteenth century.[7] They have portrayed the social movement against child labor as a response to a growing evil and have attributed the decline of youth employment after 1900 to the success of Progressive Era reforms embodied in compulsory schooling and child labor legislation (Cubberley 1909; Ensign 1921; Abbott 1947; Trattner 1970; Ashby 1985; Licht 1992).

The late-nineteenth-century rise in the employment rate of youths, however, should be rather surprising. Between 1870 and 1900 real wages rose, including those of the lowest paid workers (Abbott 1905; Coombs 1926; Long 1960; Rees 1961; Williamson and Lindert 1980). Since, in the cross-section, child labor is associated with low relative incomes of fathers (Haines 1979, 309; Horan and Hargis 1991, 590-593), the rise in real wages would be expected to lead to a *lower* incidence of child labor over time. Throughout most of the nineteenth century compulsory school laws were either absent or not enforced (Tyack 1974). Nonetheless, school expenditures rose, school attendance increased, and illiteracy fell (Fishlow 1966; Bowles and Gintis 1976; Field 1978; Tyack 1974). Since children's primary alternative to paid labor is schooling, this evidence of a voluntary increase in school attendance might be supposed to imply greater costs and lower returns to child labor, again leading to a reduction in its incidence (Solmon 1970; Lindert 1978; Landes and Solmon 1972; Tyack 1976; Rodgers and Tyack 1982).

Edwards' finding of a rising incidence in the employment of children also clashes with the findings of demographic and social historians that, during the last half of the nineteenth century, fertility declined and parental attitudes toward children shifted from opportunistic to altruistic. According to this view, parents chose to have smaller numbers of children and began to downplay their children's value as economic assets and as potential contributors to family income. Increasingly children came to be viewed as precious (priceless) beings upon which parents were willing to lavish considerable affection, education, and protective attention (Lindert 1978; Becker 1981; Zelizer 1985). Presumably, these precious children would not face an *increasing* risk of work.[8]

The traditional story about the employment behavior of older men is that their involvement in paid labor declined continuously over time. The standard estimates were developed by Durand and are displayed in Table 1. According to Durand, "The decline in the proportion of men remaining in the labor force at these ages (65 and older) was pronounced and continuous during the whole period from 1890 to 1940" (Durand 1948, 34). Indeed, his estimates indicate a drop from 68.2 to 63.2 percent in the labor

force participation rate of males 65 years of age and older between 1890 and 1900 and a further decline to 55.6 percent by 1920. Durand offers two explanations for this decline. First, over time, in response to the rise in incomes and the spread of corporate-based pension plans, more men decided to "retire voluntarily to live on savings or pensions after the age of 60 or 65" (Durand 1948, 34). Second, this economy-wide tendency was reinforced by the shift of employment out of agriculture and into industry. In industry, especially in large corporate firms, according to Durand, older workers experienced an increasing difficulty in retaining their jobs (Durand 1948, 34-35).

Ransom and Sutch have challenged this traditional view, arguing instead that overall retirement rates were stable, at least between 1870 and 1937. Their work relied on the published Census returns for 1870 and 1880 and on the Public Use Microdata Sample (PUMS) from manuscript returns for 1900 and 1910 (Ransom and Sutch 1986, 1988). Recently the Ransom and Sutch view received additional support from work by Costa (1995), Moen (1994), and Carter and Sutch (1996) who develop evidence suggesting that many farmers retired about the turn of the century and that they retired at rates very comparable to those of workers in the industrial sector. These findings raise the possibility that retirement may have been common in the nineteenth century as well.

The employment of married women is thought to trace a U-shape with industrialization (Goldin 1994). Though comprehensive statistics for the pre-industrial era in the United States are not available, documentary evidence suggests that the overwhelming majority of married women were involved in the household production of commodities for consumption and sale in that early period (Goldin 1986b). The rise of the factory system and the separation of home from work is thought to have reduced married women's involvement in the production of market goods. By 1890, when the first economy-wide figures become available, only 4.6 percent of married women appear to have been directly involved in the production of commodities (See Table 1). Improvements in women's education and the spread of white collar work over the twentieth century reversed the nineteenth-century trend away from married women's remunerative work. Married women's labor force participation rate rose to 9.0 percent by 1920 and 11.7 percent by 1930. The dislocations of World War II together with rising wage rates, improvements in women's education, and a reduction in women's fertility reinforced the positive trend toward paid work in the second half of the twentieth century (Goldin 1983, 1989, 1990, 1994).

Accuracy of the Published Occupation Returns for 1880

This paper has its origins in our discovery that certain occupational statistics which appear in the official publications of the Tenth Census differ radically from those collected by the census enumerators in 1880. The enumerators' reports can be analyzed using the Public Use Microdata Sample (PUMS) from that census. The 1880 PUMS is a random, 1-in-100 sample of the original enumerators' returns. Except for the recoding of some number of persons returned with the occupation "housekeeper" and a small number of laborers and nurses, occupation in the 1880 PUMS was coded exactly as entered by the original census enumerators. The recoded housekeepers were assigned a code that allowed them to be easily identified and the enumerators' tally replicated (Ruggles and Menard et al. 1993, 10).

The magnitude of the discrepancies between the enumerators' occupation returns and the published census occupation figures for youth, the aged, and prime-age married women are shown in Table 2. These tables display for males and females, respectively, an index calculated as the ratio of the PUMS total (multiplied by 100) to the published census figure for each of the "grand divisions of industry."

For youth, the elderly, and all women, the employment ratios displayed in Table 2 are substantially in excess of 1.00. Among males, the PUMS indicates 33 percent more employment among youths and 29 percent more employment among the aged than the published reports. Among younger, prime age, and older females, the differentials are 35, 44, and 146 percent, respectively. There are no such discrepancies between the published and PUMS totals for 1900, 1910, and 1940.[9] Instead, all of these ratios are close to 1.00. There is also no large discrepancy for prime-aged men in 1800.

The discrepancies between the published numbers and the sample values for 1880 are large and entirely outside deviations that might arise from sampling variance.[10] The published census volume indicates labor force participation rates for youthful males and females of 24.4 and 9.0, respectively while the PUMS suggest rates of 32.4 and 12.1 percent. The published census volume shows a rate of 64.3 percent for older males (60 and older) while the PUMS shows a rate of 82.8 percent. The published volume displays rates of 17.1 and 5.2 percent for women 16 to 59 and 60 and older while the PUMS implies rates of 24.5 and 12.5 percent, respectively.[11] Altogether, the PUMS implies 1.7 million more gainful workers in 1880 than does the published volume.

If the PUMS values are accurate representations of the census enumerators' gainful worker totals, the gaps of these magnitudes between the enumerators' and the published totals call into question the standard estimates for 1880 and also force us to reconsider trends in labor force participation of children, older men, and married women during the last third of the nineteenth century. This is because the 1880 published figures are the most central of the three nineteenth-century censuses providing information on occupations by age and gender. The 1890 Census uses a different set of age categories than 1870, 1880, and 1900 (Edwards 1943, Anderson 1988). Edwards dealt with this incomparability between the 1890 and surrounding censuses by estimating the youth employment rate for 1890 as the simple arithmetic mean of the 1880 and 1900 rates. Moreover, since the 1870 Census of occupations was collected and tabulated using the same protocol as that employed in 1880, our discovery of a discrepancy between the enumerated and published totals for 1880 suggests that there may be a similar discrepancy between the enumerated and published totals for 1870. Until a public use sample is collected from 1870, however, we cannot test this suggestion directly. As it is, we must form a judgement about the trends in labor force participation by age and gender in the period by comparing estimates for 1880 with those for 1900.

If PUMS figures are accepted and substituted for those from the published volumes, the stylized facts of long-term change in the employment of youth, the elderly, and married women would have to be revised. The 1880 youth employment figures in the PUMS are well above the standard estimates for 1900. Substituting the PUMS rate for the published rate presented by Edwards for 1880 suggests that youth employment *fell* between 1880 and 1900. The similarity in the protocol for the 1870 and 1880 censuses suggests that 1870 published youth employment rate may also require an upward revision. If this speculation is supported, then child labor would show a steady decline throughout the period of industrialization. The 1880 employment rate for older males in the PUMS is over 25 percent higher than the published rate. If the rate from the PUMS is accepted it would provide strong support for the traditional view of Durand and would make it difficult to sustain the revisionist work of Ransom and Sutch, Costa, Moen, and Carter and Sutch. There is no published employment rate for married women in 1880, but the PUMS allows us to construct one. It is 12.3 percent. If accepted, it would call into question both the low reported rate for 1890 (only 4.6 percent) and the extreme U-shaped pattern of married women's paid employment with industrialization.

Explaining the Discrepancy Between the Published Occupation Figures and the PUMS

Which figures ought to be accepted -- those which appear in the published census volumes or those in the PUMS? The answer depends, in part, on one's beliefs about the source of the discrepancy. We examine three possibilities.

The Sampling Bias Hypothesis

The 1880 PUMS may be have been selected in an inappropriate fashion. If the PUMS sample was drawn in a way that misrepresented the gainful occupations of youth, the aged, and women this could explain the patterns we have observed.

The sampling bias hypothesis is easily rejected. The 1880 PUMS sampling procedure is well-described and, if properly executed, each dwelling, family, and individual would have had a 1-in-100 probability of inclusion in the sample (Ruggles, Menard, et al. 1993, 3 and Ruggles 1995). Our tabulations, displayed in Tables 2 and 3 suggest that the sampling procedure *was* properly performed and that the sample is, indeed, representative. In Table 2 the ratios corresponding to the gender and broad age groupings in the population are all equal to 1.00. Table 3 displays ratios of PUMS to published totals according to race, nativity, geographic region, and urban residence. The quantitatively most important groups -- Whites, Blacks, and the native-born -- appear in their proper proportions. There is an over-representation of Chinese and Japanese and an underrepresentation of Indians (Native Americans). These groups are small, however. Their lack of representativeness cannot materially affect the gainful worker counts. The most potentially damaging disproportions are geographic. Table 3 shows a slight over-representation of residents of the Old South and the Pacific states. Youth, and also older males and women, in the Old South show a higher gainful worker rate than those in other regions of the country (though youths in the Pacific states show a lower rate). Even this distortion makes little difference for the PUMS employment rates, however. Had the geographic distribution of the PUMS perfectly reflected that of the population in the published census, the youth gainful worker rates for the nation as a whole would have been 32.2 and 11.5 percent for males and females respectively. The PUMS rates for these groups are 32.4 and 11.6 while the published rates are only 24.4 and 9.0 percent. The slight geographic distortions of the PUMS make little difference for the implied employment rate of older males or married women, either.

A final piece of evidence against the sampling bias hypothesis comes from comparisons of the PUMS and published youth, aged, and female employment totals for 1900, 1910, and 1940. These reveal an almost perfect correspondence. Since these PUMS were collected according to the same sampling procedures adopted for the collection of the 1880 sample, we conclude that the discrepancy between the 1880 PUMS and published employment totals cannot be regarded as an artifact of PUMS construction. Instead, the discrepancy appears to reflect some action in the Census Office after the collection of the enumerators' manuscripts and prior to the publication of the tabulated totals.

The Mistabulation Hypothesis

Perhaps clerical errors in the Census Office can explain the divergence between the enumerators' manuscripts and the published totals in 1880. Daniel Scott Smith, who drew his own sample from the 1880 census manuscripts for a study of old age, has suggested this possibility.[12] Smith conjectured that when tallying the occupational results into a three-column worksheet (corresponding to the three broad age groupings, 10-15, 16-59, and 60 and older) with many individual occupations listed, the tallier would focus on finding the correct occupational row upon which to enter the tally mark. Once the occupation line was located, over 85 percent of the cases would fall into the broad age category 16 to 59 and would be properly recorded in the middle column. Since youthful and older laborers were comparatively rare, there might have been a tendency to misrecord some of these workers erroneously into column two. If Smith's conjecture is correct it would explain the apparent underreporting of both youthful and older workers in 1880.[13] It would also imply that youthful and older workers were underreported in 1870 since the tabulation forms and procedures were the same. Finally, it would explain the consistent behavior of the participation rates of older men across the two censuses when examined by state and specific occupation by Ransom and Sutch (1989, 175-183).[14]

It is difficult to reject the Mistabulation Hypothesis outright since there are other puzzling discrepancies between the PUMS and the published occupation tallies. To take one example, the number of youthful males reported with an occupation in Alabama is over twice as great as in neighboring Mississippi (64,918 verses 32,330) despite the fact that the economies, population, and racial composition of these states are quite similar.[15]

If we attribute the discrepancy between the PUMS and the published youth and aged employment totals to tabulation errors, then the choice of figures to use in the long-run employment series is clear. We ought to substitute the PUMS figures although that would still leave teh under-reporting of prime-aged women's occupations unresolved. Before embracing this strategy, however, we consider a third possibility.

The Editing Hypothesis

Our Editing Hypothesis states that Census officials in Washington decided to remove a large number of youthful, older, and female workers prior to the tabulation and publication of the gainful worker count. One difficulty with this hypothesis is that, to the best of our knowledge, there is no record of any such action in the annals of the *Tenth Census*.[16] Another difficulty with our Editing Hypothesis is that it indicts the Superintendent of the 1880 Census with knowingly, willfully, and secretly directing the alteration of the enumerators' returns. The Superintendent of the 1880 Census was no faceless bureaucrat, either. He was Francis Amasa Walker (1840-1897), one of the most prominent economists and statisticians of the late-nineteenth century. Walker's accomplishments include presidency of the Massachusetts Institute of Technology (1881-1897), the American Statistical Association (1883-1897), and the American Economic Association (1885-1892). He was the first to develop explicitly the concept of "perfect competition" (Newton 1967, 24) and was the author of over a hundred technical books and articles on economics and statistics. Walker, who also superintended the Census of 1870, is generally credited as the first Bureau Chief to approach the job in a professional manner. He insisted on examinations for the Bureau clerical staff, wrested control over appointment of enumerators away from local politicians, and developed the general methodology still in use today by agencies the world over that collect and present mass data (Newton 1967, 150; Magnuson and King 1995, 28-29). The Editing Hypothesis would appear to accuse this seemingly scrupulous, public-spirited, and highly-decorated man with frivolous or perhaps malicious tampering of the public record. To help motivate such an audacious hypothesis we turn to an examination of Walker's character and career.

The Character and Career of Francis Amasa Walker

Francis Amasa Walker, son of the prominent antebellum political economist Amasa Walker, became convinced early in life of the importance of reliable statistical information for effective public policy formation. After military service during the Civil War, Walker took up teaching and journalism and was soon drawn into a national debate on the forthcoming federal census of 1870 (Munroe 1923, 3-29, 101-117). In an influential article published in the *Atlantic Monthly* in 1869, Walker argued that a dramatic change from the 1860 census protocol was required to achieve an accurate statistical portrait of the "industry of the country." In Walker's view, "not two thirds, certainly, of the national production is represented in these pretentious tables" into a logical scheme accurately reflecting the different branches of industry (Walker 1869, 689).

Manufacturing statistics were bound to be misleading indicators of the output and structure of the economy, in Walker's view, since "the products of the industry of many classes must, at the best, escape direct enumeration" (Walker 1869, 698). Occupational statistics might be used to circumvent this problem, Walker felt, though those collected in 1860 were not very serviceable.

> It is absolutely impossible from these tables to construct anything like a satisfactory scheme of the actual distribution of the people among the different branches of industry (Walker 1869, 699).

Walker attributed most of the deficiencies of the 1860 census to insufficient editing of occupation returns by the Washington office. In 1860, occupations were often tabulated under a variety of headings.

> Thus, under the head "Domestics," we have the following entries; Alabama, none; Arkansas, 797; California, none; Connecticut, none; Delaware, 1688; Florida, 631; Georgia, none; Illinois, none; Indiana, none; Iowa, -- more civilized, -- 358; Kansas, none; Kentucky, 1782. These States alone report: Louisiana; Maine, Massachusetts, Michigan, etc. being apparently destitute of such accessaries -- after the fact -- of civilization. Turning, however, to the title of "Servants," we find the deficiencies explained, the number reported under that head reaching the very pretty total of 559,908; New York, which had not a domestic, employing, it seems, 155,288 servants (Walker 1869, 699).

In Walker's view, the authorities had failed in their "great responsibility of combining and reducing (the occupational categories)" (Walker 1869, 699).

Walker's critique and constructive suggestions were so persuasive and influential that he was named Superintendent of the 1870 Census. Upon his appointment, Walker worked quickly to implement many of his proposals. He was not able to change the political appointment of enumerators -- the Marshal system -- for the 1870 census, but he attempted to compensate for inadequacies and inconsistencies in enumeration by developing detailed instructions for enumerators and tripling the size of the Washington office where the manuscript returns were corrected, tabulated, and published. He also authorized resurveys where Census Office findings were challenged (Anderson 1988, 88-89; Magnuson and King 1995, 28-29). Walker pioneered innovations in format and presentation such as four-color maps and new classification systems designed to make census statistics more understandable and useful (Anderson 1988, 92). In the words of an admirer:

> There is a mania in many figures, as there is madness in much learning; but Mr. Walker is strong enough to resist the contagion, and doubles the value of his work by the buoys and lighthouses erected on its seas of numerals (Quoted in Munroe 1923, 114).

In the *Ninth Census'* Table of Occupations these buoys and lighthouses appeared as an organization of the detailed data according to the "grand divisions of industry" -- Agriculture; Professional and Personal Service; Trade and Transportation; and Manufacturing, Mechanical, and Mining. Walker also added demographic detail, reporting the number of persons engaged according to gender and three broad age categories -- 10 to 15, 16 to 59, and 60 years and upwards -- as another interpretative aid. The result of this herculean effort of imagination, purpose, and effective control is widely regarded as our nation's earliest reliable marker of the occupations of the people. When Walker was reappointed to direct the 1880 census he preserved the 1870 occupational categories and used a much-enhanced budget to display his findings in even more detailed, varied, and innovative forms.[17] In Edwards' influential "Comparative Occupational Statistics in the United States" (1943), Walker's occupation statistics for 1870 and 1880 form the starting point of the analysis.

Evidence of Walker's Editing of Census Occupation Returns

There is ample evidence -- both direct and indirect -- that, under Walker's direction, Census Office officials in 1870 and 1880 carefully examined and in some cases altered manuscript occupation returns before their tabulation and publication. Indeed, since Walker criticized his predecessor for failing

in his "great responsibility of combining and reducing" occupation categories, we would expect him to approach the manuscript figures as an intermediate product which needed scrutiny and possible revision before their release to the public.

One straightforward admission that enumerators' returns were altered before tabulation and publication, is in a discussion of an unexpectedly small increase between 1870 and 1880 in the number of individuals who reported an agricultural occupation. The "deficit" in agricultural occupations, the Census speculated, may have been due to a misreporting of agricultural laborers as simply "laborers" in the manuscript returns.

> In some parts of the country, where agriculture is in a high degree predominant, there is doubtless a certain tendency to drop the qualifying adjective and speak of 'laborers' simply. *Wherever it has been within the power of the Census Office to apply a correction to the error it has been done,* yet there probably remains a certain amount of fallacious classification resulting from the failure of enumerators duly to characterize this class of persons (U.S. Census Office 1883, 703-704. Emphasis added.).

A second admission of editing appears in a note stating that only "gainful and reputable occupations" were tabulated (U.S. Census Office 1883, 704). This explains why the 74 "prostitutes" and 12 "gamblers" included in the PUMS do not appear in the published occupational tabulations.[18]

The apparent goal of this editing was to produce a more accurate statistical portrait of the "industry of the country." "Laborers, Not Specified" were reclassified as "Agricultural laborers" wherever the Census felt such workers were occupied in farming. Prostitutes and gamblers were omitted from the occupational tally not because of "any sensitiveness at the mention or recognition of these classes as actually existing in the community." Rather, Walker wanted to avoid misrepresenting the case.

> (T)he numbers thus reported must be wholly inadequate to the fact, and that a seeming count of them in the census would have the effect to *mislead rather than to instruct* (U.S. Census Office 1872, 659. Emphasis added.).

We have been unable to uncover any direct evidence of editing of the occupation returns of youths, the aged, and women. On the contrary, all of the remarks about the official employment figures of these groups are reassurances that the published tabulations are exactly the expected ones. In its discussion of youth employment, the Census comments that the difference between the youth population

and those with occupations "is substantially equal to the number of children attending school who do not, through any considerable portion of the year, pursue any gainful avocation" (U.S. Census Office 1883, 704). In regard to the employment of the elderly the Census states:

> The explanation of the number of persons 60 years of age and upward returned without occupation is so manifest as not to require to be even alluded to (U.S. Census Office 1883, 704).

There is no hint as to why 35.7 percent of older males returned without occupation was more "manifest" than the 17.2 percent implied by the enumerators' returns. In its discussion of women's employment the Census declares that the large number listed without a gainful occupation could be easily explained by the "wives, mothers, or grown daughters, keeping house for their families or living at home without any special avocation" (U.S. Census Office 1883, 704). If anything, the Census Office suggests that its employment figures for youth and women are *too low*.

> (W)omen and children employed in factories are omitted in large numbers.... Especially in the case of children, who are part of the time in school and part of the time in the mill, would the mother or father be likely to give a preference to the former over the latter mode of statement, particularly when, ... the child comes within the period of life respecting which there is *a legal requirement* of school attendance (U.S. Census Office 1883, 709).

There is some indirect evidence that the occupation returns of youth, the aged, and women were edited, however. In the published returns not even one youth -- male or female -- between the ages of 10 and 15 is listed under the occupational headings "Farmer and Planter," "Lawyer," "Manufacturer and Officials," "Physicians and Surgeons," and "Traders and Dealers." In the 1880 PUMS 739 boys and 56 girls report one or another of these occupations. In the published returns 22 occupations list *no* women aged 16 to 59. In the PUMS 63 women report one of these occupations.[19]

More significant, perhaps, is the discovery by the developers of the 1880 PUMS -- Ruggles, Menard and their colleagues -- that many women who were returned by the enumerators with the occupation "Housekeeper" were apparently removed from the gainful occupation totals by census tabulators. When Ruggles, Menard et al. interpreted the occupational response "housekeeper" according to 1880 enumerator instructions and coded it as "domestic" they found they had 80 percent more domestic servants than the 1880 published tabulations despite the fact that "this was one of the largest occupational

categories to start with." They decided to recode to a non-occupational response all those women who gave their occupation as "housekeeper" and who were related to the head of the household since, "(C)oded (as the gainful occupation housekeeper/domestic) the level of married women's employment in 1880 was much higher than other sources imply and suggested a great temporary upward spike in the historical trend of married women's work (Ruggles, Menard, et al. 1993, 9). In Table 2 we included these housekeepers in the gainful occupation total. Our reinclusion of housekeepers boosted the ratio of enumerated to published occupation totals for young women from 1.30 to 1.35; for women 16 to 59 from 1.05 to 1.44; and for women 60 and older from 1.50 to 2.46. It seems clear that Walker authorized quite a bit of editing of the enumerators' occupation returns before their publication.

We believe that the evidence warrants the following conclusions:

1) Walker was a man of integrity, imagination, and sophistication, dedicated to developing occupation statistics which would accurately portray the industrial structure of the nation's economy.

2) Walker was concerned with the internal consistency of the occupation tally as well as with its overall level.

3) Walker relied upon a wide range of evidence to form prior expectations about the level of occupational attachment and its distribution across age and gender groups. He considered it his *duty* to edit enumerators' reports whenever he felt they were inconsistent with other evidence.

In addition, we have formulated two working hypotheses:

1) To achieve a consistent measure of *attachment* to gainful occupations, we suggest that Walker questioned the reported occupation of all those he suspected of part-time employment. We suspect that this would lead him to question the labor force attachment, in particular, of youths attending school; women of any age living with spouses, parents, or children; and the elderly with some visible, nonwage means of support.

2) We further suspect that, to achieve consistency across similar categories, Walker edited out positive responses which implied that the person had broken the law. This might include, in addition to occupations such as prostitution and gambling, the illegal work of minors and women.

Tests of the Editing Hypothesis

Here we test our two working hypotheses, exploring the extent to which they are able to identify *systematic* patterns of editing of the manuscript returns. Our tests rely on the PUMS together with published tables on gainful employment by state, gender, and age group available in the *Statistics of the Population of the United States at the Tenth Census* (1883).

Hypotheses Relating to the Occupations of Youths

In his discussion of the measurement of the occupations of youth, Walker expressed concern over two particular problems. One was youths' tenuous attachment to paid labor. The second was their potentially deceitful parents. Walker was presumably also alert to problems of "unlikely" occupations and of consistency in the reporting of youth occupations across states with and without compulsory schooling and child labor laws. Here we attempt to test whether these factors can account for state-by-state differences between enumerators' returns and the published gainful worker totals for youths.

Our test relies on weighted least squares regression analysis. The dependent variable in our analysis is the difference between the enumerated (multiplied by 100) and published youth occupation totals for each state for which the enumerated total exceeds the published.[20] Observations are weighted by the reciprocal of the enumerated youth population. Our independent variables are proxies for labor force attachment, parents' incentives to deceive enumerators, and Walker's efforts to achieve internal consistency. We measure weak labor force attachment by the number of youths 10 to 14 who were simultaneously attending school. As we have already noted, the Census view was that (many) "children attending school ... do not, through any considerable portion of the year, pursue any gainful avocation" (U.S. Census Office 1883, 704). We partition the simultaneous school attendance measure into two independent variables; one indicating the numbers of young workers simultaneously attending school in states with a compulsory schooling law in 1879, the second indicating the number in states without such laws.[21] The estimated coefficient on each of these variables indicates the fraction of scholar-workers Walker removed before publishing the occupation totals. A coefficient equal to one implies that all young scholar-workers were removed while a coefficient of zero suggests that none were. A difference between the coefficients on the young-scholar-worker variable for states with and without compulsory schooling laws would lend support to the view that Walker's editing protocol varied systematically according to

whether a compulsory schooling law was in place, with the relative magnitudes of the coefficients indicating where Walker felt the greatest editing efforts were needed to achieve consistency.

A second set of independent variables attempts to assess the impact of child labor laws on editing. Since these laws applied only to young workers employed in manufacturing, mining, and the trades, we have computed the number of workers aged 10 to 14 in these three sectors. To measure the impact of child labor laws we also partition our manufacturing-mining-trade measure into two variables, one indicating the number of youthful workers in these sectors in states with a child labor law in 1879, the second indicating the number in states without such a law. As in the case of the young-scholar-worker measure, the coefficient on each term indicates the proportion of workers edited out of the published occupation totals; a difference between the coefficients measures the difference in editing between states with and without laws.

Finally, we include indicators of occupations that Walker may have eliminated from the gainful occupation count altogether. One is the number of youths whose enumerated occupation was considered "unlikely" by Walker, as evidenced by the fact that no youths were reported with such an occupation in the published census totals. A second is girls who returned their occupation as "Housekeeper." The presumed rationale is that these girls were not working for pay as required by the enumerators' instructions but instead were engaged in unpaid domestic production for their own families which, by convention, is excluded from the gainful occupation count. A coefficient of one on either the "Unlikely" or "Housekeeper" occupation variables indicates that youths with such occupations were deleted from the gainful occupation count. A coefficient of zero indicates that they were kept in the count but given another occupation title. The constant term indicates how well the variables in the equation can "explain" the editing. A constant term equal to zero would indicate that the variables in our equation account for the systematic editing of the enumerators' returns.

Separate regressions were run for males and females. Results, reported in Table 4, are consistent with the Editing Hypothesis. The coefficients on the scholar-worker and manufacturing-mining-trade terms in the equations for boys and girls are all positive, suggesting that youths with these characteristics were edited from the occupation rolls. The coefficients are not measured precisely, however, so that it is impossible to tell whether Walker edited differently in states with and without laws. We therefore

estimated a second set of equations in which the number of scholar-workers and the number of youths employed in manufacturing, mining, and trade are indicated without regard to the existence of a state law. In both of these equations the worker-scholar and manufacturing-mining-trade terms are positive and significantly different from zero. We can reject the hypothesis that none of these youths were edited out. The point estimates suggest that 40 percent of male and 87 percent of female worker-scholars and 70 percent of the male and 41 percent of the female manufacturing, mining, and retail workers under the age of 15 were removed from the occupation rolls before publication. The point estimates of the coefficients on the "Housekeeper" variables are large. We cannot reject the hypothesis that Walker ordered *all* youths reporting this occupation removed from the gainful occupation rolls. The coefficients on the "Unlikely" occupations measure are also large, but are measured so imprecisely that we cannot conclude that they are different from zero. If we accept zero as the coefficient, it would suggest that youths whose occupation was reported as, say, "Farmer" were assigned another occupation, perhaps farm laborer, and retained in the occupation count. On the other hand, if we accept the coefficients as estimated, it would suggest that over half of these individuals were edited out of the occupational statistics. Finally, the constant terms in both equations are not statistically significantly different from zero.

In sum, our weighted least-squares regression results are consistent with a restricted version of the Editing Hypothesis. There appears to have been a systematic effort to remove the young (aged 10 to 14) who were either enrolled in school or engaged in manufacturing, mining, and trade from the gainful occupation count. Youths recorded with the occupation "Housekeeper" also appear to have been removed. This editing of young workers, however, appears to have been directed equally at all youth, regardless of whether the state in which they lived had a compulsory schooling law or child labor law. When we remove from the enumerators' occupation totals the 40 percent of male and 87 percent of female workers 10-14 who were also in school and 70 percent of male and 41 percent of the female youths 10 to 14 in manufacturing, mining, and trade, plus all of the youths recorded as "Housekeepers" as indicated by the point estimates of our equations, we reduce the gainful employment rate of youthful males from 32.4 to 26.0 percent and the rate of youthful females from 12.1 to 9.2 percent. The published rates for these groups are 24.4 and 9.0 percent, respectively.

Hypotheses Relating to the Occupations of the Aged

Since Walker's only explicit comment in 1880 about the occupational returns of the aged was that the published returns needed no explanation, we develop the details of our editing hypotheses from Walker's general concern to eliminate marginal workers. In 1870, the Census explained the low published participation rate of older men by the "number of persons retired from active pursuits by reason of an acquired competence, of support secured from grown children, or of advanced age" (U.S. Census Office 1872, 798). Unfortunately the 1880 Census returns provide no easy way to discern who might have sufficient wealth to support themselves. We can, however, distinguish older men living with their grown children and those of highly advanced ages. We also conjecture that Walker might have ordered the removal of men who reported an occupation but who also reported some unemployment. For males we included an additional term to test whether editors might have removed those who reported themselves as the household head but who had a younger, employed male relative living with them. For older women we included a measure of whether she indicated an occupation of "Housekeeper". As in the case of youth, the dependent variable is the difference between the enumerated (multiplied by 100) and published occupation returns for those states for which this value is positive. The regressions were weighted by the reciprocal of the enumerated group population. The interpretation of the coefficients is the same as in the case of youths.

The results, reported in the first two columns of Table 5, provide additional support for the Editing Hypothesis. For older males and females, those reporting some unemployment appear to have been removed from the published figures. Males who are not household heads themselves but who live in the household of a relative and females who report the occupation "Housekeeper" were also apparently edited from the published totals. We can reject the hypothesis that these coefficients are zero and cannot reject the hypothesis that they are equal to one. Census tabulators apparently did not challenge the employment status of older male household heads, even when a younger, employed male relative was present. The coefficient on this measure is not significantly different from zero. Age per se appears to have prompted the removal of approximately 11 percent of women over 70, but not of any men.

As in the equations for youths, the constant term in both equations is not statistically significantly different from zero, implying that we cannot reject the hypothesis that the variables in our equations

adequately account for the difference between the enumerators' returns and the published occupation totals for these groups. Removing all of the "housekeepers," "unemployed," and 11 percent of the employed women 70 years of age and older closes the gap between the enumerated and published employment rate for older females. Removing all of the "relatives" and the "unemployed" from the male gainful worker count closes part, but not all of the gap. The rate for older females drops from 12.5 to 5.5 percent while that for males is reduced from 82.8 to 68.3 percent. The published gainful worker rates for these groups are 5.2 and 64.3 percent, respectively.

Hypotheses Relating to the Occupations of Prime-Age Women

Walker made no direct reference to the occupation returns of women aged 16 to 59. In the absence of any specific guidelines, we decided to test the hypothesis advanced by Ruggles, Menard, et al. that Walker removed all of those women who returned their occupation as "Housekeeper." As we have already noted, the removal of "Housekeepers" from the gainful occupation count reduces the discrepancy between the PUMS and published totals for women aged 16 to 59 from 1.44 to 1.05. What we want to test here is whether it can also explain state-to-state differences between the PUMS and published figures.

Regression results reported in the far right-hand column of Table 5 suggest that it can. The single term "Housekeeper" accounts for 74.9 percent of the variance across states in the difference between enumerated and published occupation returns for prime-age women. A second interesting result of this exercise is that the estimated coefficient on "Housekeeper" in our equation is significantly greater than one. This suggests that Walker also edited out some women who reported a legitimate gainful occupation, but who had personal characteristics similar to those of housekeepers. These results suggest that, in addition to biases of census enumerators discussed in the earlier literature, the editing of enumerators' returns in Washington, D.C. had a considerable impact on the recorded labor force participation rate of women in 1880.

Taken together, the results of this series of exercises provide substantial support for the Editing Hypothesis. Francis Amasa Walker appears to have directed that a substantial number of youthful, elderly and female workers be removed from the returns prior to their publication. The pattern of his editing is consistent with his self-proclaimed goal of developing occupation statistics appropriate for

assessing the "industry of the country."

Editing and Long-Term Labor Force Trends

While Walker was primarily interested in assessing the industry of the country, the principal use of the occupation statistics of 1880 today is in identifying trends in labor force participation over time.[22] For this modern goal, *consistency* in the definition of the labor force across census years is crucial. Is Walker's editing is appropriate for our contemporary purposes? Did Walker's editing make the published occupation statistics for 1880 more or less consistent with those for later years? We consider the employment count of youth, older males, and married women in turn.

Youth Employment

There are two paragraphs of instructions to enumerators that are relevant in assessing the comparability of the 1880 youth occupation count with that from later years: those which relate to the enumeration of the occupations of youths and those which deal with the treatment of school children. Instructions regarding youth employment were virtually identical in both years.[23] The only substantive difference is a reference in the 1900 instructions to "paragraph 162" which refers to worker-scholars and introduces a *change*. Paragraph 162 reads:

> 162. Report a student who supports himself by some occupation according to the occupation, if more time is given to that, but as a student, if more time is give to study. Thus report a student who does stenographic work as a student unless more of his time is spent in stenography. Report a salesman in a grocery store, who attends a night school as 'salesman, groceries,' because most of his day is spent in the store (Ruggles, Menard, et al. 1993, 141, para 162).

The "more time" rule was not in place in 1880. Rather, in 1880 *all* boys and girls "earning money regularly by labor, contributing to the family support, or appreciably assisting in mechanical or agricultural industry" were to be recorded with an occupation. At least one scholar has argued that this change in instructions to enumerators "greatly" reduced the number of children recorded with occupations in 1900 as compared with previous censuses (Bliss 1905).[24]

It is important to note, however, that the "more time" rule would have affected only those youthful workers attending school more than six months per year. According to figures compiled by the U.S. Office of Education, the average length of the school term in 1880 was 130.3 days or 6 months ((130.3/5)/(30.4/7)) while the average number of days attended per enrolled pupil was 81.1 days or 3.7 months (U.S. Bureau of the Census 1975, Series H521 and H522, 346). The length of the term varied widely, however. In urban areas in the North and West the school term was 169 days or 7.7 months. In rural areas the average term ran only 125 days or 5.7 months in the North and West and a mere 76 days or 3.5 months in the South (Meyer, Tyack, Nagel, and Gordon 1979, 594, Table 1). Moreover, in rural areas the school term was structured around youth work obligations. Thus, a rural school imposed even less of a constraint on youth employment opportunities than these numbers might suggest. According to Tyack:

> (A)ttendance varied enormously from day to day and season to season, depending on the weather, the need for labor at home, and the affection or terror inspired by the teacher. During the winter ... older boys attended .... During the summer ... older children worked on the farm (Tyack 1974, 16-17).

The potential impact of a "more time" rule in 1880, then, depends upon the fraction of worker-scholars who lived in urban areas and attended school for more than six months a year.

We calculated the fraction of worker-scholars living in urban areas using the 1880 PUMS. We discovered that only 2.4 percent of males and 5.1 percent of females age 10-14 and reporting both work and school lived in places with a population of more than 2,500. In other words, almost all of the worker-scholars in 1880 lived in rural areas. With short rural school terms structured to complement their work schedules, it seems probable that these scholar-workers would have been enumerated with an occupation even if the "more time" rule had been in effect in 1880.[25] To achieve consistency between the youth occupation returns of 1900, all but 2.4 percent of the male and 5.1 percent of the female worker-scholars who were removed for the tabulation ought to be retained in the youthful gainful occupation count for 1880. Since the instructions regarding youth employment were otherwise identical in 1880 and 1900, *all* of the youth in manufacturing, mining, and trade should be returned. In other words, the enumerators' returns rather than the published census figures for youth employment in 1880 ought to be accepted. Substituting the enumerators' for the published youth employment rates for 1880

and making an appropriate adjustment to the rate for 1890 implies that the trend in child labor has been steadily downward at least since 1880.[26]

The Employment of Older Men

After concluding that the PUMS data based on the original enumerations are superior to the published data for youth in 1880, it might be supposed that we would similarly prefer the PUMS data for older men. The situation here, however, is more complex. Changes in the instructions to the enumerators between 1880 and 1900 may be thought to have *increased* the number of older men who were enumerated with occupations. These changes have been described and debated at length in Ransom and Sutch (1986, 1989) and Moen (1987). Briefly, instructions applicable in 1900 but not in 1880 included the explicit instruction to include occupations for retired individuals (Ransom and Sutch 1986, 8-9). Some older men were thus recorded with occupations such as "retired lawyer," "landlord," or "capitalist." or were recorded with their "habitual occupation" but reported twelve months of unemployment.

However, as Moen has pointed out, the practical impact of these changes turned out to be small. The employment rate of men 60 and over calculated from the 1880 and 1900 PUMS rose only slightly from 82.8 percent to 83.6 percent. This suggests that there was little change in the labor force attachment of older men between 1880 and 1900. Of course, Walker's editing in 1880 lowered the published figures for that year considerably, making it appear that there was a substantial *increase* in the employment rate from 64.3 to 83.6 percent.

Ransom and Sutch have argued that the 1900 PUMS figure for older men with occupations is highly misleading precisely because it includes many older men who were actually retired but who reported occupations and at the same time reported extensive unemployment. They used the 1900 PUMS to remove men with six months or more of reported unemployment and came up with an estimate that 64.11 percent of men 60 and over were employed (and 66.06 percent were in the labor force) (Ransom and Sutch 1986, Table 1). This rate could be compared with the *published* 1880 figure if we are prepared to believe that Walker's editing was also intended to remove workers who were actually retired from active pursuits by reason of an "acquired competence, of support secured from grown children, or of advanced age." By this reckoning, the employment of older men seems to have changed little between

1880 and 1900: from 64.3 percent to 64.1 percent. The conclusions about the trend in the labor force participation of older men would be the same as those advanced by Ransom and Sutch, that is, that there was no trend. We suggest that the most meaningful estimates of the employment of older males are those estimated by Walker for 1880 (and 1870) and by Ransom and Sutch for 1900.[27]

The Employment of Married Women

Earlier we argued that Walker's apparent decision to exclude "Housekeepers" from the occupation count of women explains much of the difference between the enumerated and published returns for women aged 16 to 59. The treatment of "Housekeepers" is also crucial in our assessment of long-term trends in the labor force participation rate of married women. If "Housekeepers" are included in the gainful worker count for 1880, then 12.3 percent of married women overall would be gainfully occupied according to the PUMS. If "Housekeepers" are excluded, then the participation rate for this group is only 5.7 percent.

The evidence we have developed regarding "Housekeepers" is somewhat contradictory, but on balance leads us to believe that they should be included in the gainful worker count. Consistency in treatment from census to census clearly dictates their inclusion. "Housekeepers" are included in the gainful occupation totals for 1900 and later. In 1900, 11.0 percent of employed married women had this occupational title while 41.6 percent of employed married women were occupied in some domestic service occupation.[28]

Instructions regarding the enumeration of "Housekeepers" in 1880 and 1900 also dictate their inclusion since these instructions are virtually identical in the two years.[29] In both years enumerators were admonished to look for evidence of a wage or salary payment for the housekeeping service and to exclude women whose housekeeping was performed for their own families. No one has produced evidence that the enumerators failed in this reponsibility in either 1880 or 1900.

We conclude that the housekeepers should be retained in 1880. But can we accept the implication that as many as 6.6 percent of married women worked for pay as housekeepers in others homes in 1880 when only 0.8 percent of married women were recorded with this occupation in 1900? The work of social historians on married women's involvement in domestic service in the nineteenth century suggests

that the 6.6 percent level for 1880 is not implausible. Dudden emphasizes the importance of temporary domestic help in the nineteenth century, arguing that it was common for families to hire domestic help to assist with the arrival of a new baby, an illness, and for busy seasons such as spring and fall housecleanings and the harvest. She goes on to outline the role of married women:

> Sometimes married women performed this kind of temporary daywork. Late in the (nineteenth) century Sarah Christie Stevens, a Minnesota farm wife, hired Mrs. Vigell, the wife of her husband's hired man, to help with the spring cleaning. The Vigells lived down the road. "Mrs Vigell came up this morning, washed floor of little room, pantry and dining room, cleaned out cellar and white-washed walls -- she is coming again in the morning to wash safe and stain and scrub tables, clean up back room and help move stove" (Dudden 1983, 32).

Dudden's account suggests that married women who performed housekeeping activities for pay might not look very different from married women who did not. Katzman argues that black women, especially, had a strong preference for live-out domestic service (Housekeeping) from the earliest days, since so many employed black women were also married. "(B)y specializing in household labor and selecting live-out work," Katzman argues, "they could at least continue to run their own homes and families" (1978, 84). Reinterpreting data collected by Hill (1929) on the occupations of women, Katzman goes on to show that live-out domestic service predominated in the urban South as early as 1900 (Katzman 1978, 90). He contends that the transition of domestic service from predominantly live-in to live-out in the early part of the twentieth century was the result of the northward migration of married black women who brought with them their preference for live-out service (Katzman 1978, 88). Married women who needed cash but wanted to continue to run their own homes and care for their families were the ones who took up "Housekeeping."

If most of the married "Housekeepers" worked like Mrs. Vigell -- on an intermittent basis -- then Walker, whose goal was to derive a measure of the full-time labor force, would have been correct to exclude them from the gainful occupation total. On the other hand, consistency with figures for later years would seem to require the inclusion of these "Housekeepers." The most puzzling implication of a decision to include housekeepers in 1880 is that it would imply a very dramatic decline in the labor force participation rate of married women between 1880 and 1890 from 12.3 to only 4.6 percent followed by a small rise to 5.6 percent in 1900. For now we say that 12.3 percent seems to us to be the right

number for 1880 and leave the problem of reconciliation with figures for later years as a topic for future research.

Editing and the Sociology of Nineteenth-Century American Official Statistics

The official statistics of the United States were often described in the nineteenth-century as the best in the world. They were more voluminous than those of any other country. They were also extraordinarily accessible, objective, and reliable. Americans devised institutional safeguards shielding statistical agencies from meddling by politicians and interest groups. They developed questionnaires, instructions to enumerators, classification systems, coding instructions and an administrative structure that greatly increased uniformity and precision.[30] The objectivity and accuracy of the resulting data made the American protocols a model for official statistical agencies around the world.[31]

Much of the credit for the high quality of official statistics belongs to Francis Amasa Walker and his protege, Carroll Davidson Wright.[32] Walker and Wright elevated the nineteenth-century American "celebration of numbers" that has been so well-described by Patricia Cline Cohen to its greatest height. In Cohen's view, nineteenth-century Americans valued numbers because:

> (T)hey marked the progress of the era, and...they were reputedly objective and precise and hence tantamount to truth. Numerical facts trounced opinions and were supposed to foster community consensus, because all thinking people would naturally agree if they possessed total and accurate information (Cohen 1982: 225).

In a telling passage, written near the end of his long and productive career, Wright expressed in moving terms his faith in the role such statistics might play in describing social relationships.

> Bright colors mixed and laid by an artist's hand on canvas might not tell at the close of another century of the work of our generation. Glowing words of description by gifted writers might not mean all the interpretation we give it in our time. But cold, enduring Arabic characters will remain through the ages to come, as truthful as of old. They are the fitting and never-changing symbols by which to tell the story of our present state, so that, when the age we live in becomes the past of successive generations, the story and the pictures shall be found to exist in all the just proportions in which it has been set by ourselves (Wright 1908:15).[33]

The particular genius of Walker and Wright was to translate the popular love of and confidence in statistics into a vehicle for addressing the social upheavals that accompanied industrialization. If workers and bosses, the native- and foreign-born, or parents and children were at odds, then a thoughtfully-constructed statistical survey might suggest a resolution. In his introduction to the *Eight Annual Report* of the Massachusetts Bureau of Labor in 1877, Wright expressed his belief that statistical information was the key to resolving social problems:

> Any means which the Legislature can adopt which will add to the information of the people on subjects which concern their daily lives are of untold value....To popularize statistics, to put them before the masses in a way which shall attract, and yet not deceive, is a work every government which cares for its future stability should encourage and enlarge (Quoted in Goldberg and Moye 1985: 8).

By and large, Walker and Wright were content to let others to draw conclusions from the data they collected. This was because they felt the implications of properly collected and presented statistics would be obvious to everyone. In his presidential address to the American Social Science Association in 1886, Wright made this point explicitly: "With the enlightenment of the workers of society, the reforms so much sought for will come as a *natural consequence*" (Quoted in Goldberg and Moye 1985: 8. Italics added). In the view of Walker and Wright, the statistician -- not the social scientist or politician -- held the key to the resolution of political and social conflict. Another reason for the statistician to let others draw conclusions from the data was that such a distanced stance made it easier for the statistician to be scrupulously honest and objective. On this point, Wright was fond of quoting the advice given to him by Walker upon his assumption of the leadership of the Massachusetts Bureau of Labor Statistics at the beginning of his career:

> (Y)our office has only to prove itself alike superior to partisan dictation and to the seductions of theory in order to command the cordial support of the press and of the body of citizens....(T)he difficulty of collecting statistics in a new country requires much indulgence; and I have strong hopes that you will so distinctly and decisively disconnect the Massachusetts Bureau of Labor Statistics from politics, from dependence on organizations, whether of workingmen or of employers, and from the support of economical theories, individual views, or class interests -- as to command the moral support of the whole body of citizens and to receive the co-operation of all men of all occupations and of all degrees, without reference, however, either to their degrees or their occupations (Wright 1908: 11-12).

Both Walker and Wright made scrupulous honesty the hallmark of their careers. They instituted reforms to insulate their statistical agencies from political influence and devised protocols to insure that the data they collected was accurate, unambiguous, and accessible to the public. Walker's innovations in professionalizing the appointment of and developing detailed instructions for enumerators, improving the speed and accuracy of the tabulation, and making the results more accessible to the general public have already been noted. Wright, whose office specialized in surveys, developed new methods of collecting representative samples and developed a variety of techniques for gaining the confidence of respondents (Carter, Ransom, and Sutch 1991). Throughout his career Wright was aware of the possibility of sabotage of the process by an official interested in "the support of economical theories, individual views, or class interests." Yet, after thirty-two years of work with federal and state statistical agencies, Wright concluded that, to an astonishing degree, the state labor bureau chiefs had indeed taken an objective approach to their work. He insisted that:

> (N)o matter for what reason they were appointed, no matter how inexperienced in the work of investigation and of compilation and presentation of statistical matter, no matter from what party they came, and whether in sympathy with capital or with labor, and even if holding fairly radical socialistic views--the men have, almost without exception, at once comprehended the sacredness of the duty assigned them, and have served the public faithfully and honestly, being content to collect and publish facts without regard to their individual bias or their individual political sentiments. As soon as a man realizes that he is giving to the world a fact, he feels the necessity of accuracy, and that to distort the information collected would be to commit a crime worse than any ordinary lying, because it would mislead legislators and others and fix a falsehood in the history of the State (Wright 1901b: 1-2).[34]

To modern ears, tinged with cynicism about bureaucratic respect for accuracy and truth, these comments might seem incredulous. Yet scholars have concluded that Wright's comments seem to have accurately expressed the prevailing ethic of his time and guided the collection and presentation of official American statistics in late-nineteenth century America.

Furthermore, scholars have interpreted Walker's and Wright's statements to mean that, except where they corrected obvious errors and inconsistencies, the census and other government statistical agencies published the exact responses to questions asked. Thus to develop consistent statistical series scholars have scrutinized conceptual shifts (from the gainful worker to the labor force concept, for example) or looked for changes in the completeness of canvasses or in instructions to enumerators. A

good example is Alba Edwards' (1943) approach to the occupation statistics for 1870 through 1940. Possible changes in the fidelity of the published totals to the enumerators' reports have not even been considered.

Our discovery that Walker apparently authorized an extensive and secret editing of the manuscript occupation returns is therefore a surprise. Walker had the opportunity to describe his revisions. As we have shown, he did in fact describe other editing he had authorized. Thus there are two puzzles: why did he do it and why didn't he tell?

It is easy to imagine why he did it. His purpose in collecting occupation statistics was to assess the industrial structure of the country. The inclusion of marginal workers would distort this picture. In retrospect, he might have preferred to have issued enumerator instructions that eliminated these marginal workers from the gainful worker count in the first place. Given that they were recorded as gainfully occupied, however, it was better to edit them out of the published totals than to slavishly include them. This approach, which we associate today with post-modernism, holds that a number is not a "fact" but a text. As text, numbers require interpretation. Some evidence that Walker embraced such a view was offered by Carroll Wright in a passage contrasting the "statistical mechanic" to the "real statistician":

> (T)he real statistician, the man who is working out the process of making (statistics),...puts an integrity and devotion into his work that is not surpassed in any other line of official conduct. The spirit of the modern statistician lies in the precepts laid down by General Walker and in the fact that there is something deeper and more comprehensive than the mere statement of figures, for the statistician must have the spirit of what...Mr. North has called ethical philosophy (Wright 1908, 13-14).[35]

Our suggestion, then, is that Walker removed marginal workers so that the occupational count would more accurately portray the industrial structure of the country.

Why didn't he tell? We don't know. Perhaps he was afraid of igniting social protest. Even the publication of his lower, modified figures brought a storm of protest over the evils of child labor from labor and social reform groups (Abbott 1908, 36). He may have worried that the unadjusted figures would exacerbate rather than ameliorate social problems.[36] Perhaps he was afraid of undermining public trust in the integrity and reliability of official statistics. He may have worried that the public would perceive his actions as an effort at manipulation or distortion.[37] Perhaps he was afraid of being held up

to public scorn. He was appointed Superintendent of the Census after ridiculing his predecessor. He may have worried that a public disclosure that his instructions to enumerators produced a badly-distorted picture of full-time labor input would expose him to the same sort of judgement.[38]

Conclusions

The release of the public use microdata sample drawn from the population enumeration of 1880 provides an opportunity to evaluate long-standing beliefs about the extent and trend in the labor force involvement of youths, older workers, and married women in the United States. It also provides insights into the social construction of official statistics in that era. Our examination of this newly-available source leads us to conclude that the enumerators' occupational returns from the federal census of 1880 were heavily edited prior to their publication. If we are correct, then our discovery has two important implications.

Regarding labor market trends, our study suggests that the incidence of child labor was far greater in 1880 (and quite probably in 1870) than has been previously thought. The required correction is large enough to reverse the supposed upward trend between 1880 and 1900. Our findings further suggest stability in the employment trend of older males about the turn of the century and a need to reexamine the view that a deep U-shaped trend characterized the involvement of married women over the period of industrialization. These revisions to the accepted view raise important questions about the role of compulsory schooling and child labor legislation, Social Security and private pensions, and the appearance of white collar jobs in accounting for the evolution of the American labor force over time.

Regarding the sociology of official statistics, our study suggests that American statisticians may have taken less of a positivist approach to their work than previously thought. If true, this fact is important both for data users and for those interested in the social construction of official information.

NOTES

Susan B. Carter is Professor of Economics, University of California, Riverside, CA 92521. Richard Sutch is Professor of Economics and History, and Director of the Institute of Business and Economic Research, University of California, Berkeley, CA 94720. He is also a Research Associate of the National Bureau of Economic Research.

Versions of this paper were presented to the Social Science History Association, the Stanford-Berkeley Demography Seminar, and the All-University of California Group in Economic History Conference. We thank members of these audiences, especially Matthew Sobek, Carolyn Moehling, and Daniel Scott Smith for constructive comments and suggestions. Though not a member of one of these audiences, Margo Anderson also offered valuable advice and counsel. Some of the data utilized in this paper were made available by the Inter-university Consortium for Political and Social Research. The data for the *Census of Population, 1880 United States: Public Use Sample (1 in 100 Sample) Computer File* were originally collected by Steven Ruggles and Russell R. Menard of the University of Minnesota, Department of History, Social Science Research Laboratory. The data for the *Census of Population, 1900* were originally collected by the Center for Studies in Demography and Ecology, University of Washington. We gratefully acknowledge the work of Roger Ransom, Tom Tang, Charles Wetherell, and Hongcheng Zhao all of the University of California, Riverside for assistance in handling this data set. We thank the National Science Foundation, the National Institute for Aging, the Institute for Business and Economic Research, University of California, and the Academic Senate of the University of California for financial support at various stages of this research.

1. In 1820 and 1840 the census occupational inquiry was limited to information on the number of persons in each household engaged in major *industrial* categories such as agriculture, commerce, and manufactures. In 1850 more occupational information was obtained, but from free males 16 years and older only. In 1860 the occupational inquiry was extended to free women 16 years and older. No occupational statistics were collected in 1830.See Edwards (1943, 87).

2. While Walker also directed the 1870 Census, it is thought to have seriously underenumerated the population (and their occupations) in 13 Southern states.

3. The published occupation figures for 1890 are considered to be of high quality, but are difficult to compare with figures for earlier and later years since they use a different age grouping than those for 1870, 1880, or 1900.

4. Scholars have identified three types of bias in official U.S. statistics, especially in statistics pertaining to the employment of women. *Instructions to enumerators* defining "gainful occupation" so as to exclude the unpaid work of women in the home is one. A second is the *behavior of enumerators* who may have assumed that women would not work outside the home and thus failed to ask about their occupations (and those of children). A third is the surrepitious *behavior of tabulators* in the punch-card tabulation era (1890 to 1940) who may have deliberatedly miscoded "unusual" occupational responses in an effort to avoid the scrutiny of their supervisors. See Abbott (1924), Jaffe (1956), Smuts (1959, 1960), Conk (1981), Ciancanelli (1983), Goldin (1986a), Bose (1987), Folbre and Abel (1989), and Anderson (1992). The interpretation of "gainful worker" has also been debated, in particular the correspondence between "gainful worker" and the modern "labor force" concept in the case of older males. See Achenbaum (1978), Smith (1978), Ransom and Sutch (1986, 1989), Moen (1987), and Margo (1993).

5. On the political importance of the guarantee of accuracy in published official statistics in nineteenth century America see Wright (1877, 1908); Waldin (1911); Leiby (1960); Cohen (1982, 1990); Goldberg and Moye (1985); and Carter, Ransom, and Sutch (1991).

6. Several recent analyses also emphasize the heavy use made of child labor in early American manufacturing establishments relative to agriculture. See Field (1978) and Goldin and Sokoloff (1982, 1984).

7. Not all historians have accepted this view. Lebergott, for example, asserted that:

> (C)hild labor never became a really substantial source of labor in America -- not

> even in farming or homework. As early as 1833 it was on the wane in the cotton textile industry, where child labor had probably been more important than in any other industry (Lebergott 1964, 126).

Lebergott did not join the debate over the impact of industrialization on the employment of children, however.

8. Of course it is not difficult to imagine several plausible explanations for a rise in child labor despite these progressive developments. Perhaps industrialization did create a new kind of light factory job that children could perform, much as Abbott (1908) and others have suggested (See also, U.S. Census Office 1883, 707 and Spargo 1906). The increasing importance of new immigrants from southern and eastern Europe who may not have shared the "modern" American values about the inappropriateness of putting children to work before age 14 or 15 may have been the proximate cause of the rise. Perhaps it is simply incorrect to postulate a rise of parental altruism in this period. For example, Parsons and Goldin (1989) report finding evidence of opportunistic behavior by parents during this era. None of these suggestions, however, have been sufficiently explored, either singly or in combination, to make a convincing case.

9. There is no PUMS from the 1870 Census and the enumerators' manuscripts from the 1890 Census were destroyed in a fire.

10. For the 1880 PUMS dwellings rather than individuals were sampled. The calculation of standard errors for such cluster samples is complicated, as Ruggles (1995, 40) notes. Since the number of individuals within any single demographic group in a single household is likely to be small, the impact of clustering on the standard errors of our index are small. In any case, the number of individuals age 10 and older in the PUMS exceeds 350,000. With a sample this large, standard errors are extremely small.

11. The discrepancy between the published figures and the manuscripts in the labor force participation rates of males 16 to 59 is small. The rates for this group in the published source

is 93.4 while the rate implied by the manuscripts is 92.8 percent.

12. The sample collected by Smith was conducted for the Newberry Library (Smith 1978). A codebook provides more information (Jensen et al. 1985). Smith's hypothesis was originally described in a letter to Ransom and Sutch.

13. It would not account for the underreporting of female workers, however.

14. Margo Conk (1981) has developed a different "mistabulation hypothesis" for the occupation statistics in the censuses of 1890 through 1940 when the Hollerith tabulating system was used. Since the Hollerith system reduced the cost of identifying certain kinds of responses, Census Bureau officials began trying to identify errors by pulling *all* unusual occupation codes for reexamination and comparison with the original enumerators' manuscripts. See Truesdell (1965, 89-103). Conk suggests that clerks working under such a system would feel subtle pressure to produce a "reasonable" tally and avoid the scrutiny of their supervisor. These pressures might lead them, in Conk's view, to deliberately mistabulate the enumerators' records, biasing the published occupation figures to show fewer women in "men's" occupations and fewer children at work. Since the Hollerith system was not used in 1880, Conk's hypothesis is not directly relevant to our argument. Moreover, some informal experiments lead us to believe that the subtle pressures on tabulators described by Conk probably had only a small quantitative impact on the published occupation totals in 1900. Our informal experiments were conducted with the 1900 sample of the enumerators' manuscripts from the Census of 1900 (1900 PUMS) when the Hollerith system was in use. Comparisons between the PUMS and published figures for female and youth employment and the distribution of female employment across occupations reveal a close correspondence between the PUMS and the published occupation totals. Since pressures on clerks to avoid the scrutiny of supervisors would have been smaller in 1880 before the Hollerith system was in operation, it seems unlikely that Conk's version of the mistabulation hypothesis could explain the very large discrepancy between the enumerators' and published occupational figures we have discovered in the 1880 Census.

15. In 1880 the population of Alabama exceeded that of Mississippi by 12 percent while the male youth employment figure was more than twice as large.

16. There is no suggestion of such editing in the standard histories of the U.S. Census (Wright 1900; Holt 1929; Anderson 1988). Edwards' careful *Comparative Occupation Statistics for the United States, 1870 to 1940* (1943) is also silent on this issue. Of course, since Edwards worked only with the published record, he could not be expected to have been aware of any discrepancy between the enumerators' returns and the published figures or to have suspected editing. King and Magnuson (1993, 84-85) describe "Provisions for Correcting Final Returns," but focus exclusively on procedures designed to identify under- and over-counts. We made two visits to the National Archives in an effort to uncover evidence of editing. We conducted a thorough examination of all items in Record Group 29, "Records of the Bureau of the Census." We also inspected Record Group 48, "Records of the Office of the Secretary of the Interior," Item #284, "Records Relating to the Ninth and Tenth Decennial Censuses," and Item #286, "Records Relating to the Eleventh Decennial Census." We found no evidence that the published returns were edited.

17. By the 1880 census Walker had won his bid for more and smaller enumeration districts and the substitution of professional for politically-appointed census takers. According to all accounts, the quality of the 1880 enumeration was vastly superior to that in any previous census. The quality differential was especially notable for hard-to-enumerate categories. See Anderson (1988) and King and Magnuson (1993).

18. There is also evidence of very careful inspection if not actual editing. See, for example, the extended discussion of "Anomalous Entries" (U.S. Census Office 1883, 711).

19. The excluded occupations with the most numerous responses in the PUMS are "Gardeners, nurserymen, and vine-growers," "Draymen, hackmen, and teamsters," "Carpenters and joiners," and "Rope and cordage makers." There are also a number of occupations reported by older women in the PUMS which were not reported in the published occupational tabulations.

20. The published total exceeded the enumerated in eight states for boys and in 11 states for girls. The overwhelming majority of states for which the published totals are greater than the enumerated are in the South. We conjecture that the enumerators' returns were supplemented to correct a perceived undercount. For boys, Alabama shows a published rate of 71.5 percent but an enumerated rate of only 58.4 percent. For girls, the published rate of 34.5 percent in Louisiana far exceeds the enumerated rate of only 21.2 percent. For older males in many Southern states the published also exceeds the enumerated rate.

21. Our coding of states according to the existence of compulsory schooling and child labor laws relies upon a classification developed by Landes and Solmon (1972, 56-57, Table 1). We coded compulsory schooling laws for all states that had adopted what Landes and Solmon term an "Education Law" by 1879. According to Landes and Solmon, most of these laws "required all children between 8 and 14 years of age to attend school for at least 12 weeks per year. Exceptions of children lacking mental and physical ability, and for the poverty of the parent or guardian" (Landes and Solmon 1972, 58). We coded child labor laws for all states for which Landes and Solmon marked a child labor law in existence in 1879. Landes and Solmon's definition of a state with a child labor law is one which "prohibits a child from accepting employment in retailing, manufacturing, or mining unless he has attained a minimum age" (Landes and Solmon 1972, 58).

22. This is not the only use of these statistics, of course. Kuznets (1952), Gallman (1975), Weiss (1991), and Geib-Gundersen and Zahrt (1995) have used these figures to measure labor inputs into the production process. This use is precisely the one that inspired Walker.

23. In 1880 enumerators were instructed:

> Neither will the doing of domestic errands or family chores out of school be considered an occupation. "At home" or "attending school" will be the best entry in a majority of cases. But if a boy or girl, whatever the age, is earning money regularly by labor, contributing to the family support, or appreciably assisting in mechanical or agricultural industry, the occupation should be stated (Ruggles, Menard, et al. 1993, 136).

In 1900 the instructions read:

> The doing of domestic errands or family chores out of school hours where a child regularly attends school, is not an occupation. But if a boy or girl, above 10 years of age, is earning money regularly by labor, contributing to the family support, or appreciably assisting in mechanical or agricultural industry, the kind of work performed should be stated (see paragraph 162) (Ruggles, Menard, et al. 1993, 146, para 219).

24. Referring specifically to paragraph 162, W.L. Bliss argued:

> (T)he instructions to enumerators for the present census (of 1900) were materially different from the instructions for the two previous censuses, and ... this difference must have resulted in greatly reducing the number of children reported as employed in gainful occupations (Bliss 1905 p. 257).

He appeared to be unaware of the editing of enumerators' returns prior to their publication in 1880 or of the small number of days of school attended by rural students in 1880.

25. By 1900, the average length of the school term had increased to 144.3 days or 6.6 months and the average number of days attended soared to 99.0 days or 4.6 months -- a 25 percent increase. Moreover, bureaucratic structures began to dominate informal community school organizations, even in rural areas (Tyack 1974). It seems plausible that an increasing fraction of youthful workers were employed only after school and during holidays. The "more time" rule may have been instituted to eliminate a growing corps of students with "summer jobs" from the occupation rolls.

26. Our suggestion that the trend during the last one-third of the nineteenth century was *away* from the use of child labor is given support by an independent set of census tabulations -- the Manufacturing Census reports on employment of wage workers in "manufacturing." Beginning in 1870 the inquiries made to firms regarding employees asked for the "average number of hands employed" over the preceding year, divided among men, women, and children. Data from the four manufacturing censuses spanning 1870 and 1900 show that the percentages of total manufacturing employment represented by children fell from 5.6 to 3.2 percent. See U.S.

Census Office, 1902.

27. The situation in 1910 is different since in that year new instructions were issued to prevent the recording of occupations for the retired. See Margo (1993) and Carter and Sutch (1994).

28. We classify as "Housekeepers" in 1900 those with occupation codes 156 ("Housekeepers and Stewards") and 947 ("Hskp & Stwd, Dom Resident"). Domestic service workers include the "Housekeepers" plus those with occupation codes 127 ("Nurses (not specified)"), 131 ("Servants"), 148 ("Domestic Service -- not specified"), and 121, ("Laundry work (hand)").

29. In 1880 the instructions read:

> The term "housekeeper" will be reserved for such persons as receive distinct wages or salary for the service. Women keeping house for their own families or for themselves, without any other gainful occupation, will be entered as "keeping house." Grown daughters assisting them will be reported without occupation.

In 1900 instructions state:

> Return as a *housekeeper* a woman who receives a stated wage or salary for her services, and do not confuse her with a woman who keeps house for her own family or for herself, without any gainful occupation, or with a grown daughter who assists in the household duties without pay. A wife or daughter who simply keeps house for her own family should not be returned as a housekeeper in any case. (para. 184)

30. For a sociological analysis of the production of official statistics see Starr (1987, 30-46).

31. For evidence on the evaluation and international impact of late-nineteenth century U.S. government statistical protocols see Wright (1901a, b), Hanger (1904), Norwood and Early (1984), and Goldberg and Moye (1985).

32. Carroll Davidson Wright was the father of the state labor bureau movement and served as the first United States Commissioner of Labor from 1885 to 1905. Wright shared Walker's belief in the value of statistics and his commitment to objectivity in their collection, and his dedication to their accessibility. Leiby's biography of Wright (1960) describes his work in its

social, intellectual, and political context. For a more recent assessment see Carter, Ransom, and Sutch (1991). Wright paid homage to Walker on many occasions, calling him his teacher and mentor. See for example Wright (1908).

33. We have deleted a few phrases from Wright's original prose to enhance its readability.

34. William Peters (1987: 84) claims that Wright was the originator of the phrase, "figures won't lie, but liars will figure."

35. The "Mr. North" to whom Wright refers is Simon Norton Dexter North (1848-1924), director of the United States Census from 1903 to 1909. Prior to North's appointment as census director he authored a volume of the *Tenth Census of 1880* (*History and Present Condition of the Newspaper and Periodical Press of the United States*), wrote a report for the *Eleventh Census of 1890* ("Wool Manufacture"), and served as chief statistician for manufactures for the *Twelfth Census of 1900.*

36. This line of speculation does not account for Walker's failure to disclose his editing of the occupation statistics of older men, however.

37. For a discussion of the falsification of official statistics in other countries in an effort to avert political upheaval see Starr (1987, 38-9).

38. We admit to at least one problem with this line of speculation. If Walker concluded that his instructions to enumerators were deficient, then one might have expected him to adjust them after the 1870 census. In fact, instructions regarding the enumeration of the gainfully occupied are virtually identical in the two years. One possibility is that Walker's improvements in the enumeration process had a disproportionate effect on the inclusion of marginal workers so that the 1880 manuscript returns showed a disproportionate increase in the gainful occupation rate of youths, women, and older persons. To further investigate this possibility would require a public use microdata sample from the manuscript returns of 1870.

REFERENCES

Abbott, E. 1905. The wages of unskilled labor. *Journal of Political Economy* 13(3): 321-67.

_____. 1908. A study of the early history of child labor in America. *American Journal of Sociology* 14(1): 15-37.

_____. 1924. *Women in industry: A study in American economic history*. New York: D. Appleton.

Abbott, G. 1947. *The child and the state: Selected documents with introductory notes*. Chicago: University of Chicago Press.

Achenbaum, W. A. 1978. *Old age in the new land: The American experience since 1790*. Baltimore: Johns Hopkins University Press.

Anderson, M. J. 1988. *The American census: A social history*. New Haven: Yale University Press.

_____. 1992. The history of women and the history of statistics. *Journal of Women's History* 4(1): 14-36.

Ashby, L. 1985. *Saving the waifs: Reformers and dependent children, 1890-1917*. Philadelphia: Temple University Press.

Becker, G. S. 1981. *A treatise on the family*. Cambridge: Harvard University Press.

Bliss, H.L. 1905. Census statistics of child labor. *Journal of Political Economy* 13(2): 245-57.

Bose, C. E. 1987. Devaluing women's work: The undercount of women's employment in 1900 and 1980. In *Hidden aspects of women's work*, edited by C. E. Bose et al., 95-115. New York: Praeger.

Bowles, S. and H. Gintis. 1976. *Schooling in capitalist America: Educational reform and the contradictions of economic life*. New York: Basic Books.

Carroll, M. R. 1923. *Labor and politics: The attitude of the American Federation of Labor toward legislation and politics*. Boston: Houghton Mifflin.

Carter, S., R. L. Ransom, and R. Sutch. 1991. Doing the Wright thing: Carroll Wright and the state labor statistics movement. Berkeley: University of California Historical Labor Statistics Project Working Paper.

Carter, S. B. and R. Sutch. 1994. Economic activity of older men: A reconsideration of the 1900 data. University of California *Working Paper in the History of Retirement*.

_____. 1996. Myth of the industrial scrap heap: A revisionist view of turn-of-the-century American retirement. *Journal of Economic History* 56(1): *forthcoming*.

Center for Studies in Demography and Ecology, University of Washington. 1980. *United States census data, 1900: Public use sample.* Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research.

Ciancanelli, P. 1983. Women's transition to wage labor: A critique of labor force statistics and reestimation of the labor force participation of married women in the United States, 1900-1930. Ph.D. diss., New School for Social Research.

Cohen, P. C. 1982. *A calculating people: The spread of numeracy in early America.* Chicago: University of Chicago Press.

_____. 1990. The early (and soon broken) marriage of history and statistics, Boston 1839. *Historical Methods.* 23(2): 63-71.

Conk, M. A. 1981. Accuracy, efficiency and bias: The interpretation of women's work in the U.S. census of occupations, 1890-1940. *Historical Methods* 14(2): 65-72.

Coombs, W. 1926. *The wages of unskilled labor in the manufacturing industries in the United States, 1890-1924.* New York: Columbia University Press.

Costa, D. L. 1995. Agricultural decline and the secular rise in male retirement rates. *Explorations in Economic History (forthcoming).*

Cubberley, E. P. 1909. *Changing conceptions of education.* Boston: Houghton Mifflin.

Dudden, F. E. 1983. *Serving women: Household service in nineteenth-century America.* Middletown, Conn.: Wesleyan University Press.

Durand, J. D. 1948. *The labor force in the United States 1890-1960.* New York: Social Science Research Council.

Edwards, A. M. 1943. *Comparative occupation statistics for the United States, 1870 to 1940.* Washington: Government Printing Office.

Ensign, F. C. 1921. *Compulsory school attendance and child labor.* Iowa City, Iowa: Athens Press.

Field, A. 1978. Sectoral shift in antebellum Massachusetts: A reconsideration. *Explorations in Economic History* 15(2): 146-171.

Fishlow, A. 1966. Levels of nineteenth-century investment in education. *Journal of Economic History* 26(4): 418-24.

Folbre, N. and M. Abel. 1989. Women's work and women's households: Gender bias in the U.S. census. *Social Research* 56(3): 545-70.

Gallman, R. 1975. The agricultural sector and the pace of economic growth: U.S. experience in the 19th century. In *Essays in nineteenth century economic history: The Old Northwest*, edited by D. C. Klingman and R. K. Vedder, 35-76. Athens: Ohio University Press.

Geib-Gundersen, L. and E. Zahrt. 1995. A new look at U.S. productivity growth, 1880-1900: What the IPUMS reveals. Department of Economics, UC Riverside *Working Paper 95-7.*

Goldberg, J. P. and W. T. Moye. 1985. *The first hundred years of the Bureau of Labor Statistics.* Washington, D.C.: Government Printing Office.

Goldin, C. 1983. The changing economic role of women: A quantitative approach. *Journal of Interdisciplinary History* 41(4): 707-33.

_____. 1986a. The female labor force and American economic growth, 1890-1980. In *Long-term factors in American economic growth*, edited by S. Engerman and R. Gallman, 557-594. Chicago: University of Chicago Press.

_____. 1986b. The economic status of women in the early Republic: Quantitative evidence. *Journal of Interdisciplinary History* 16(3): 375-404.

_____. 1989. Life-cycle participation of married women: Historical evidence and implications. *Journal of Labor Economics* 7(1): 20-47.

_____. 1990. *Understanding the gender gap: An economic history of American women.* New York: Oxford University Press.

_____. 1994. The U-shaped female labor force function in economic development and economic history. Cambridge, Mass.: *NBER Working Paper* No. 4707.

_____ and K. Sokoloff, 1982. Women, children and industrialization in the early Republic: Evidence from the manufacturing censuses. *Journal of Economic History* 42(4): 741-74.

_____. 1984. The relative productivity hypothesis of industrialization. *Quarterly Journal of Economics* 99(3): 461-88.

Haines, M. R. 1979. Industrial work and the family life cycle, 1889-1890. *Research in Economic History* 4: 289-356.

Hanger, G.W.W. 1904. Bureaus of labor statistics in the United States. United States Department of Labor, Bureau of Labor Statistics *Bulletin 54*. Washington: Government Printing Office.

Hill, J. A. 1929. *Women in gainful occupations: 1870 to 1920*. Census Monograph IX. Washington, D.C., Government Printing Office.

Holt, W. S. 1929. *The Bureau of the Census: Its history, activities and organization*. Washington: The Brookings Institution.

Horan, P. M., and P. G. Hargis. 1991. Children's work and schooling in the late nineteenth-century family economy. *American Sociological Review* 56(5): 583-596.

Jaffe, A. J. 1956. Trends in the participation of women in the working force. *Monthly Labor Review* 79(5): 559-565.

Jensen, R. et al. 1985. *Old age in the United States, 1880*. Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research #8427.

Katzman, D. M. 1978. *Seven days a week: Women and domestic service in industrializing America*. New York: Oxford University Press.

King, M. L., and D. L. Magnuson. 1993. A procedural history of the enumeration process in the 1880 U.S. census. In *1880 public use microdata sample United States Census of Population* edited by Steven R. Ruggles and Russell R. Menard *et.al.*, Appendix B. Minneapolis, Minnesota: Social History Research Laboratory, Department of History, University of Minnesota.

Kuznets, S. 1952. *Income and wealth of the United States*. Cambridge, England: Bowes and Bowes.

Landes, W. M., and L. C. Solmon. 1972. Compulsory schooling legislation: An economic analysis of law and social change in the nineteenth century. *Journal of Economic History* 32(1): 54-91.

Lebergott, S. 1964. *Manpower in economic growth: The American record since 1800*. New York: McGraw-Hill.

Leiby, J. 1960. *Carroll Wright and labor reform: The origin of labor statistics*. Cambridge: Harvard University Press.

Licht, W. 1992. *Getting work: Philadelphia, 1840-1950*. Cambridge: Harvard University Press.

Lindert, P. H. 1978. *Fertility and scarcity in America*. Princeton: Princeton University Press.

Long, C. D. 1960. *Wages and earnings in the United States, 1860-1890.* Princeton: Princeton University Press.

Magnuson, D. L., and M. L. King. 1995. Comparability of the public use microdata samples: Enumeration procedures. *Historical Methods* 28(1): 27-32.

Margo, R. A. 1993. The labor force participation of older Americans in 1900: Further results. *Explorations in Economic History* 30 (4): 409-423.

Meyer, J. W., D. Tyack, J. Nagel, and A. Gordon. 1979. Public education as nation-building in America: Enrollments and bureaucratization in the American states, 1870-1930. *American Journal of Sociology* 85(3): 591-613.

Moen, J. R. 1987. The labor of older Americans: A comment. *Journal of Economic History* 47(3): 761-67.

_____. 1994. Rural nonfarm households: Leaving the farm and the retirement of older men, 1860-1980. *Social Science History* 18(1): 55-75.

Munroe, J. P. 1923. *A life of Francis Amasa Walker.* New York: Henry Holt.

Newton, B. 1968. *The economics of Francis Amasa Walker: American economics in transition.* New York: Augustus M. Kelley.

Norwood, J. L. and J. F. Early. 1984. A century of methodological progress at the U.S. Bureau of Labor Statistics. *Journal of the American Statistical Association* 79(388): 748-61.

Parsons, D. O., and C. Goldin. 1989. Parental altruism and self-interest: Child labor among late nineteenth-century American families. *Economic Inquiry* 27(4): 637-659.

Peters, W. S. 1987. *Counting for something: Statistical principles and personalities.* New York: Springer-Verlag.

Ransom, R. L., and R. Sutch. 1986. The labor of older Americans: Retirement of men on and off the job, 1870-1937. *Journal of Economic History* 46(1): 1-30.

_____. 1988. The decline of retirement in the years before Social Security: U.S. retirement patterns, 1870-1940. In *Issues in contemporary retirement*, edited by R. Ricardo-Campbell and E. P. Lazear, 3-26. Stanford: Hoover Institution Press.

Rees, A. 1961. *Real wages in manufacturing, 1890-1914.* Princeton: Princeton University Press.

Rodgers, D. and D. Tyack. 1982. Work, youth, and schooling: Mapping critical research areas. In *Work, youth, and schooling: Historical perspectives on vocationalism in American education*, edited by H. Kantor and D. Tyack, 270-294. Stanford: Stanford University Press.

Ruggles, S. 1995. Sample designs and sampling errors. *Historical Methods* 28(1): 40-46.

_____, J. D. Hacker, and M. Sobek. 1995. Overview: Order out of chaos: The integrated public use microdata series. *Historical Methods* 28(1): 33-39.

_____, and R. R. Menard *et al.* 1993. *1880 public use microdata sample United States census of population codebook and data file*. Minneapolis, Minnesota: Social History Research Laboratory, Department of History, University of Minnesota.

Solmon, L. 1970. Opportunity costs and models of schooling in the nineteenth century. *Southern Economic Journal* 37(1): 66-83.

Smith, D. S. 1978. Community-based sample of the older population from the 1880 and 1900 United States manuscript censuses. *Historical Methods* 11(2): 67-74.

Smuts, R. 1959. *Women and work in America*. New York: Schocken.

_____. 1960. The Female labor force: A case study in the interpretation of historical statistics. *Journal of the American Statistical Association* 55(289): 71-79.

Spargo, J. 1906. *The bitter cry of the children*. New York: Macmillan.

Starr, P. 1987. The sociology of official statistics. In *The politics of numbers*, edited by W. Alonso and P. Starr, 7-58. New York: Russell Sage Foundation.

Trattner, W. I. 1970. *Crusade for the children: A history of the National Child Labor Committee and child labor reform in America*. Chicago: Quadrangle Books.

Truesdell, L. 1965. *The development of punch card tabulation in the Bureau of the Census, 1890-1940.* Washington, D.C.: Government Printing Office.

Tyack, D. B. 1974. *The one best system: A history of American urban education*. Cambridge: Harvard University Press.

_____. 1976. Ways of seeing: An essay on the history of compulsory schooling. *Harvard Education Review* 46(3): 355-389.

United States Bureau of the Census. 1975. *Historical statistics of the United States, colonial times to 1970, bicentennial edition*. Washington, D.C.: Government Printing Office.

United States Bureau of Labor. 1910. *Report on condition of woman and child wage-earners. VI: The beginnings of child labor legislation in certain states.* Senate Doc. No. 645 (61st Cong., 2d sess.) Washington, D.C.: Government Printing Office.

United States Census Office. 1872. *Statistics of the population of the United States from the original returns of the ninth census (June 1, 1870)* Washington, D.C.,: Government Printing Office.

_____. 1883. *Statistics of the population of the United States at the tenth census (June 1, 1880).* Washington, D.C.,: Government Printing Office.

_____. 1902. *Twelfth census of the United States, taken in the year 1900. Manufactures, Volume VII, Part I.* Washington, D.C: Government Printing Office.

Wadlin, H. G. 1911. *Carroll Davidson Wright: A memorial.* Massachusetts Bureau of Statistics. *Fortieth Annual Report on the Statistics of Labor.* Boston: Wright & Potter Printing Co.

Walker, F. A. 1869. American industry in the census. *Atlantic Monthly* 24(146): 689-701.

Weiss, T. 1991. Long term changes in U.S. agricultural output per worker, 1800 to 1900. *Economic Research Working Paper Series on Historical Factors in Long Run Growth*, 23. Cambridge: National Bureau of Economic Research.

Williamson, J. G., and P. H. Lindert. 1980. *American inequality: A macroeconomic history.* New York: Academic Press.

Wright, C. D. 1877. *Massachusetts eighth annual report on the statistics of labor.* Boston: Wright & Potter Printing Co., State Printers.

_____. 1900. *History and growth of the U.S. census.* Washington, D.C.: Government Printing Office.

_____. 1901a. The working of the Department of Labor. *United States Department of Labor Exhibit, Pan-American Exposition, 1901.* Washington, D.C.: Government Printing Office.

_____. 1901b. The value and influence of labor statistics. *United States Department of Labor Exhibit, Pan-American Exposition, 1901.* Washington, D.C.: Government Printing Office.

_____. 1908. Address of Carroll D. Wright, President of the American Statistical Association, at its annual meeting in Boston, Jan. 17, 1908. *Journal of the American Statistical Association* New Series, 81: 1-15.

Zelizer, V. A. 1985. *Pricing the priceless child: The changing social value of children.* New York: Basic Books.

TABLE 1
Trends in the Employment of Youths, Older Men, and Married Women, Standard Estimates 1870 to 1930

Year	Youths 10 to 15 Years		Males 65 and Older	Married Women 15 and Older
	Males	Females		
1870	19.3	7.0	--	--
1880	24.4	9.0	--	--
1890	25.9	10.0	68.2	4.6
1900	26.1	10.2	63.2	5.6
1910	21.7	8.1	--	--
1920	16.8	5.8	55.6	9.0
1930	6.4	2.9	54.0	11.7

Sources: Edwards (1943, 97); Durand (1948, 208); Goldin (1990, 17).

TABLE 2
Occupations in 1880, Males and Females Ten and Older
Public Use Microdata Sample (PUMS) Relative to Published Census Index

Age	10 to 15	16 to 59	60 and Older	Total
		Males		
Population, 10 and Older	1.00	1.00	1.00	1.00
All Occupations*	1.33	0.99	1.29	1.03
Agriculture	1.31	1.01	1.16	1.05
Professional and Personal Service	1.22	0.96	1.45	1.00
Trade and Transportation	1.69	0.99	1.72	1.02
Manufacturing, Mechanical, and Mining	1.53	0.99	1.52	1.03
No Occupation	0.90	1.09	0.48	0.89
		Females		
Population, 10 and Older	1.00	1.00	1.01	1.00
All Occupations*	1.35	1.44	2.46	1.45
Agriculture	1.09	0.86	1.25	0.93
Professional and Personal Service	1.67	1.77	3.38	1.81
Trade and Transportation	1.77	1.65	2.08	1.67
Manufacturing, Mechanical, and Mining	1.37	1.14	1.59	1.04
No Occupation	0.97	0.91	0.93	0.92

Notes: * Includes some gainfully occupied with no particular occupation specified.

Sources: The published figures were taken from the Department of the Interior, Census Office (1883, Table XXX, 714). The enumerations were calculated from the 1880 manuscript census returns coded and documented by Ruggles, Menard, *et al.* (1993).

TABLE 3
PUMS Totals and Ratio to Published Totals, Selected Characteristics by Gender, 1880

	Males		Females	
	PUMS Number	Ratio PUMS/Published	PUMS Number	Ratio PUMS/Published
Population	255,951	1.00	246,496	1.00
Race				
Whites	221,490	1.00	212,583	1.00
Blacks and Mulattos	32,796	1.00	33,274	1.00
Indians	277	0.82	267	0.82
Chinese and Japanese	1,121	1.11	100	2.09
Nativity				
Native	218,537	1.00	215,233	1.00
Foreign	37,414	1.03	31,263	1.03
Geographic Division				
New England	19,499	1.00	20,472	1.00
Northeast	51,678	0.99	52,433	0.99
Midwest	69,608	0.99	65,269	0.99
Old South	50,412	1.03	50,147	1.03
New South	41,036	1.00	40,144	1.00
West	16,868	1.00	13,485	1.00
Pacific	6,850	1.03	4,546	1.02
Urban Residence				
Resident of one of the 50 largest cities	37,830	0.99	39,210	0.99

Sources: See Table 2.

TABLE 4
Determinants of the Editing of Enumerators' Occupation Returns, Youths, 1880
Weighted Least Squares Regressions
Dependent Variable: Difference Between Workers Enumerated and Published

	Males		Females	
Independent Variables	A	B	A	B
Constant	132.364 (184.507)	157.4436 (183.3874)	64.0756 (202.4163)	61.2948 (195.2072)
Workers 10 to 14 Attending School in States with a Compulsory School Law	.4253* (.1784)	--	.7332 (.5262)	--
Workers 10 to 14 Attending School in States without a Compulsory School Law	.4069* (.1060)	--	.8244* (.1560)	--
Workers 10 to 14 Attending School	--	.4048* (.1057)	--	.8279* (.1493)
Workers 10 to 14 in Manufacturing, Mining, and Trade, States with Child Labor Laws	.6042* (.1745)	--	.4202 (.2176)	--
Workers 10 to 14 in Manufacturing, Mining, and Trade, States without Child Labor Laws	.9596* (.2554)	--	.4424 (.3013)	--
Workers 10 to 14 in Manufacturing, Mining, and Trade	--	.7158* (.1524)	--	.4040* (.1571)
Workers with "Unlikely" Occupations	.6605 (.4490)	.7118 (.4433)	.5361 (.8497)	.5269 (.8218)
Workers with Occupation "Housekeeper"	--	--	1.5077* (.6074)	1.4972* (.5858)
R^2	0.8718	0.8648	0.7843	0.7571
N	41	41	37	37

Notes: * Significantly different from zero at the 5 percent level. Standard errors in parentheses. The unit of observation is the state. Estimates obtained by weighted least squares where the weight is the inverse of the population of youths. Estimates obtained using STATA.

Sources: See Table 2.

TABLE 5
Determinants of the Editing of Enumerators' Occupation Returns
Elderly Males and Females and Prime Age Females, 1880
Weighted Least Squares Regressions,
Dependent Variable: Difference between Workers Enumerated and Published

	60 and Older		16- 59 Years
Independent Variables	Males	Females	Females
Constant	139.7771 (105.4009)	63.7400 (44.7139)	-1099.105 (960.2112)
Lives with a Relative	1.4094* (.4286)	-.0277 (.0328)	--
Reports Some Unemployment	0.9635* (.2798)	1.7981* (.8448)	--
70 Years of Age or Older	-0.0098 (.1427)	.1078* (.0357)	--
Household Head with Employed Male Relative in Home	.0962 (.1232)	--	--
Occupation "Housekeeper"	--	1.2661* (.3139)	1.785* (0.1678)
R^2	0.9219	.9274	.7490
N	42	42	40

Notes: * Significantly different from zero at the 5 percent level. Standard errors in parentheses. The unit of observation is the state. "Lives with a Relative" means that the individual is not a household head but is related to the head. Estimates obtained by weighted least squares where the weight is the reciprocal of the relevant population. Estimates obtained using STATA.

Sources: See Table 2.