Volume Title: Studies in Income and Wealth

Volume Author/Editor: Conference on Research in Income and Wealth.

Volume Publisher: NBER

Volume ISBN: 0-870-14166-X

Volume URL: http://www.nber.org/books/unkn49-1

Publication Date: 1949

Chapter Title: The Changing Industrial Distribution of Gainful Workers: Comments on the Decennial Statistics, 1820-1940

Chapter Author: Solomon Fabricant

Chapter URL: http://www.nber.org/chapters/c5706

Chapter pages in book: (p. 1 - 45)

Part I

The Changing Industrial Distribution of Gainful Workers: Comments on the Decennial Statistics, 1820–1940

Solomon Fabricant
New York University
and National Bureau of Economic Research

This paper was prepared in a study of trends in the service industries, undertaken at the National Bureau of Economic Research with the aid of a grant from the Maurice and Laura Falk Foundation of Pittsburgh. Thanks are due also to Daniel Carson, W. Braddock Hickman, and other friends for comments, and to Robert E. Lipsey and Belle C. Nathan for aid in checking the manuscript.

Changes in the Industrial Composition of Manpower since the Civil War

Daniel Carson
Railroad Retirement Board

The author is indebted to Solomon Fabricant for reading the draft and for his many suggestions and to Winifred E. Summer for her assistance and suggestions in revising the estimates.

CHANGES AND CONTRASTS in the industrial structures of countries interest most social scientists. Industrial structure they take to be a significant characteristic of a society, worth study if not for itself then for revealing the forces and conditions bearing on other, more important, matters. The concept is prominent in the description of economic development in general, and in particular in the discussion of such episodes as the industrial revolution; it appears in many policies prescribed to further the wealth of nations; it plays a role in interpretations of political currents. All uses involve the application of an estimate of the importance assumed by different industries in a nation's economy.[1] Industrial structure is a concept of relative proportions.

In view of its significance for social science, and its inherently quantitative content, students of national income and wealth have naturally given the notion a good deal of attention. They have tried to provide figures on carefully defined aspects of industrial structures, for various countries and various periods. In the main they have devoted themselves to the aspects that are their own particular concern, namely, net value added and wealth; but a good deal of supplementary information on the industrial distribution of employment, wages, and gross value of product, for example, is also a part of their stock in trade. However, even members of this fraternity cannot make bricks without clay, though they have been known to get along without straw. At most, then, their efforts have yielded reliable figures covering limited and relatively recent periods. It is no reflection on the heroic efforts of King and Martin to state that for the United States, with which we are immediately concerned, there are really no better figures on the changing industrial structure during the 19th century, if not during the first decade or two of the 20th also, than the decennial data on gainful workers collected by the Bureau of the Census and put into shape by Whelpton, Edwards, and Carson.[2]

[1] The concept may profitably be enlarged, of course, to embrace interrelations among industries. The *tableau économique* set up in statistical form by Wassily Leontief is an example.

[2] P. K. Whelpton, 'Occupational Groups in the United States, 1820–1920', *Journal of the American Statistical Association*, Sept. 1926; A. M. Edwards, *Comparative*

Though they may be better than other statistics, how good are they? The presentation of Carson's revised estimates and an outline of his methods may properly be made the occasion for some critical remarks on the decennial statistics of the industrial distribution of gainful workers in the United States beginning with 1820. These comments may be of some interest also to those concerned with the statistics of other countries and with international comparisons.

A complete critique of a body of data is of course a huge, and in a sense endless, task. It is huge because to be thorough it must include many operations: examination of the schedules used and instructions given to respondents, enumerators, and editors; in the light of their capacity to comprehend and effectively cooperate, of the methods utilized in the field, and of the principles and categories underlying the summarization of the data; analysis of the internal consistency of the data collected; and comparison of these data, or of derivatives of them, with the data—quantitative and qualitative—collected or calculated by other methods or entirely different approaches. The task is endless because no final conclusion can be drawn concerning the adequacy of a body of data except as it is applied to some specific problem or theory. Data adequate for one purpose may be quite inadequate for another; and the number of purposes is infinite. What one can attempt, therefore, in this sort of commentary is simply to list some of the points, obvious and otherwise, that anyone using occupational statistics for the United States must bear in mind if he is not to misuse them. The reader must expect to emerge with a sense of some of the things he must consider in applying the data to his ends rather than with a definite notion of their accuracy or value for any particular purpose.

For his convenience I preface my remarks with a brief survey of the Censuses of Occupations and conclude with a summary table consolidating, with some modifications and additions, Carson's and Whelpton's figures. The readers should understand that

Occupation Statistics for the United States, 1870–1940 (Washington, D. C., 1943); and Daniel Carson, 'Changes in the Industrial Composition of Manpower since the Civil War' (see below).

many of the points noted here are referred to by Whelpton, Edwards, and Carson, as well as in the regular Census reports. I have attempted to systematize the discussion, and in some places to go into detail. The authors themselves cover some of the points more fully. My chief purpose is to provide a critical introduction to their work.

1 *The Censuses of the working force are incomplete in several respects; therefore they merely provide raw data for an approximation to an industrial distribution*

Even a simple survey of the basic census data the compilers of the industrial distribution of the working force had to use will give readers unfamiliar with the Censuses of Occupations an appreciation of the problems and difficulties.

A complete or partial census of gainful workers has been taken in the United States every ten years beginning with 1820, except in 1830. As in 1820 and 1840 workers were asked only if they were engaged in certain specified industries, not all industries were covered. In all other censuses workers were requested to report their occupation, whatever it was; and in Censuses beginning with 1910 the industry to which they were attached was also requested.

Based on these reports, an occupational classification has been published for every Census beginning with 1850, but an industrial classification for only 1820, 1840, 1910, 1930, and 1940.[3] However, even the so-called occupational classification is in fairly considerable part also an industrial classification, since many occupational categories were so defined as to be peculiar to specific industries. On the other hand, even the industrial classification, except for

[3] While information on industrial affiliation was collected in 1920, the Census published no industrial distribution for that year.

The five Census reports for 1900 and later years are 12th Census of the United States, 1900, Special Reports: *Occupations* (1904); 13th Census, 1910, *Population*, Vol. IV: *Occupation Statistics* (1914); 14th Census, 1920, *Population*, Vol. IV: *Occupations* (1923); 15th Census, 1930, *Population*, Vol. V: *General Report on Occupations* (1933); 16th Census, 1940, *Population*, Vol. III: *The Labor Force* (1943).

Censuses for years preceding 1900 are discussed, and much of the summary data collected in them reproduced, in the 1900 report, Ch. II (pp. xxix-lxiv). See also Ch. VIII and other portions of the report by A. M. Edwards, cited above.

1940 and perhaps also 1820 and 1840, is not strict since certain occupations, consisting of persons in several industries, were allocated to the industry in which most of them belonged.

Variations between Censuses have been considerable in the occupational and industrial codes utilized, in respect of both the kind of classes set up and the degree of detail. Differences in kind, even for periods as recent as 1930 and 1940, and even for apparently identical occupations, are illustrated by some of the adjustments needed to make the 1930 and 1940 occupational categories comparable (Edwards, pp. 35–48), and by the difficulties encountered by Carson. The number of occupations for which statistics were presented, 1850–1930, ranged from 584 (1860) to 218 (1890); the figure for 1940 is 451. In 1820 three industries were shown separately: agriculture, commerce, and manufactures. The 1840 Census included these plus mining, navigation of the ocean, navigation of canals, lakes, and rivers, and learned professions and engineers—seven in all. The largest number was in 1940, when 132 industries were distinguished.

Variation in detail of classes is especially troublesome when, as in Censuses before 1910, very broad occupational categories straddle many industries. Some are quite important; for example, 'laborers (not specified)'. Adjustments, frequently involving rough estimates, have had to be made for this as well as other difficulties before continuous series on a uniform classification, of either occupations or industries, could be derived from the original Census data.[4]

Besides the incomplete Censuses of 1820 and 1840, in which information on certain industries was not requested, some other enumerations were incomplete in the sense that some specific geographical area, race, or sex was inadequately covered. In the 1850 Census of Occupations, free females and all slaves were omitted. The 1860 Census included free females but continued to omit slaves. The 1870 Census failed to cover some persons in the southern states. It is hardly likely that the earlier, and perhaps also the later, Censuses, adequately covered American Indians.

[4] Considerable effort has also gone into constructing, from the Census data, series according to social-economic groupings. See Edwards, Part III, and his references to Hunt, Wright, and others, as well as to his own writings.

In addition to these changes in coverage, there were variations between Censuses in the age limit below which occupation data were not requested. Analysis of the figures suggests, however, that any error resulting is slight.[5]

Other difficulties that had to be met by adjustments and estimates arose from changes in the schedules or in the Census date,[6] or obviously incorrect or otherwise inadequate reporting. Adjustments for some of these have been made or indicated by the Census authorities: for 1890, upward, for children 10–15 in agricultural pursuits; for 1910, downward, mainly for women and children in agricultural pursuits; for 1920, upward, for the same class of workers as in 1910; for 1930, downward, for the net difference between 'omitted entries' and retired or disabled workers; and for 1940, upward, for the sum of omitted entries and the misclassification of public emergency workers.[7] In addition, the 1930 and 1940 figures have been specially adjusted to enhance their com-

[5] No age limit was specified in 1820 or 1840 though the 1820 instructions direct the exclusion of infants and superannuated persons. In 1850 and 1860 children under 16 were specifically excluded. No lower limit was specified in the 1870 schedules; however, instructions to enumerators stipulated that infants or children too young to take any part in production were to be omitted. The Bureau of the Census assumed this to mean that in effect the returns were confined to persons 10 years of age and over (1900 report, p. xxxi). This age limit was specified in succeeding Censuses through 1930. In 1940 the lower limit was put at 14 years. As the proportion of children in the working force has declined, especially since 1900, an upward bias in its reported growth may be expected. The bias is of course not eliminated though it is lessened by the overlap device. However, it is slight. In 1900, for example, about 8 percent of children 10 years old were gainful workers; the percentage for the whole 10–15 group was about 18, and for those 16–24, about 60. Almost any reasonable curve fitted to these points would suggest a percentage for 9 year old children of less than 6, and perhaps an average of no more than 2 or 3 percent for the age group 5–9. This would mean less than 1 percent of all reported gainful workers in 1900. In 1930 children of 10–15 constituted about 1.4 percent of the reported total. Even if the 1940 percentage were half that, the error in the 1930–40 comparison due to the neglect of the under 14 age group would be much less than 0.7 percent. As most of these children were engaged in agriculture (Edwards, p. 97; the percentage of working children aged 10–15 engaged in agricultural pursuits was 70 in both 1930 and 1870), the error for that industry would be greater.

[6] The Census date ('as of' which the figures are given) has usually been June 1; but on occasion it has been January 1, April 1, April 15, August 7, or the week of March 24–30.

[7] Daniel Carson makes an additional adjustment in the 1890 figures for children over 15 years of age. For 1910 both Carson and Clarence Long make a rather greater adjustment than the Census. Both question also its adjustment for 1920.

parability, mainly by putting the 1930 gainful workers data on a labor force basis, but also by adjusting the 1940 figures for National Youth Administration student workers.

Under (or over) coverage has thus had to be estimated, and continuous occupational and industrial classifications of workers built up. The largest part of the job has been done by the Bureau of the Census itself, particularly in Edwards' valuable report. Whelpton's main contributions were to fill in the gaps for 1820–60 (his results were accepted by the Bureau of the Census); to make preliminary adjustments of the aggregates after 1860 (since superseded by the Bureau's); and to prepare a rough industrial distribution for 1820–1920 (the 1870–1920 figures have been superseded by Carson's estimates). Carson's main contribution was to estimate the industrial distribution on a consistent basis for the entire period 1870–1930, using Census industrial classifications of 1910 and 1930 and Census occupational classifications of these and other years as presented in the regular decennial reports or reworked in Edwards' monograph.

2 *The industrial distributions are based on a definition of gainful work or production that for some purposes is too narrow and distorts the relative importance of and changes in certain industries*

Collation of the instructions to enumerators in the various Censuses of Occupations brings out the difficulties of satisfactorily defining 'gainful worker' and indicates vividly how these difficulties have troubled the directors of the censuses.[8] Some questions arise because the position of the line dividing persons counted among gainful workers from those excluded is essentially arbitrary, depending as it does on how broadly one defines economic production.

a) The chief group of persons affected are women (and a few men) engaged primarily or entirely in the operation of their own households, and other members of the family assisting them.

[8] Instructions for all censuses through 1890 appear in C. D. Wright and W. C. Hunt, *History and Growth of the United States Census* (Government Printing Office, 1900); for censuses 1870–1930, in the 1930 Report on Occupations, pp. 23–31; for 1940, in the 1940 Report, Appendix.

I need not repeat the well-known reasons why the exclusion of housewives and similar houseworkers is recognized as arbitrary. The Census does include housework done by outsiders for a recompense, small though it may be; unpaid family labor contributing to the family's income, including a certain amount of income in kind as well as cash in the case of farms, and perhaps also service and retail establishments; and, in Censuses beginning with 1910, work entailed in keeping boarders, if substantial.[9] Some estimators of national income explicitly include an allowance for the value of housewives' services; others provide supplementary figures to indicate the order of its magnitude; practically all warn of the paradox that ensues when a man marries his housekeeper. If one is interested in the changing industrial distribution of the population it is arguable, I think, that explicit account be taken of the shifts between home and outside work; i.e., that unpaid housework should be included as a category under domestic service. Such inclusion would have enormous influence on the relative importance of domestic service since the number of unpaid houseworkers far exceeds that of paid domestics.[10] In any case, some account must be taken of this large group in interpreting whatever

[9] In the 1910–30 Censuses the keeping of boarders or lodgers was considered a gainful occupation if the person so engaged relied upon it as his principal means of support. In the 1940 Census a housewife keeping 5 or more boarders or lodgers was specifically defined as a member of the labor force. Instructions for Censuses prior to 1910 do not mention keeping boarders.

As late as 1930 almost 10 percent of all families included one or more lodgers, and in 1940, 8 percent (16th Census, *Families*, General Characteristics, p. 4). (In both years households with more than 10 lodgers were excluded from the category of private families.) In earlier years the percentage must have been substantially higher because of the large immigration, a disproportionate fraction of which consisted of adult males. George Stigler points out that in 1901, according to the *Eighteenth Annual Report of the Commissioner of Labor*, over 20 percent of families in urban areas reported income from boarders and lodgers (see his 'Domestic Servants in the United States, 1900–1940' NBER, *Occasional Paper 24*, April, 1946, p. 29).

[10] If the 'domestic and personal service' category is divided to show domestic service separately, as we do later, how shall the keeping of lodgers be treated? On the principle of classifying two-occupation persons by their chief occupation, few housewives would be placed in the personal (excluding domestic) service group, since only 0.8 percent of all families kept more than 3 lodgers in 1930 (*Abstract of the 15th Census*, p. 411).

figures on the working force are published. This can be done only roughly, since the 1940 Census alone has collected adequate information on household work. In that year 28.9 million females 14 years and over were reported as engaged in unpaid housework and 13.0 million in the labor force; males, about a quarter million in unpaid housework. If we estimate the number of females so engaged in earlier years by assuming (as seems approximately confirmed by some figures available for 1920 and 1930)[11] that the percentage of females doing either housework or gainful work equaled the percentage of males gainfully occupied, we have the accompanying figures for 1870–1940. Since the proportion of females of

[11] The 1930 Census reported the number of homemakers, not also gainfully employed, as 24.5 million (*Abstract of the 15th Census*, p. 413). Since a homemaker is defined as the female member of the family who is responsible for the care of the home and the family, this figure fails to include other females, such as daughters, working at home without pay. The more inclusive figure, cited above, is available for 1940 only (16th Census, *Population*, II, Characteristics, Part 1, p. 12). Female 'home housekeepers' without gainful occupation, 16 years and over, are roughly estimated to number 22–23 million in 1920 (J. A. Hill, *Women in Gainful Occupations 1870 to 1920*, Census Monograph IX, 1929, pp. 5–6). Apparently this estimate covers grown daughters helping, as well as housewives.

The assumption that the fraction of females of working age who are gainfully occupied or doing housework at home is equal to the fraction of males of working age who are gainfully occupied is that used by R. G. Hurlin and M. B. Givens, in their chapter, Shifting Occupational Patterns, in *Recent Social Trends*; see p. 274, Table 1, and p. 279, note 9; cf. also Edwards, *op. cit.*, p. 90. (Hurlin and Givens apply the method to the group 16 years and older, rather than to the group 10 and older, as is done above.) The two fractions were approximately the same in 1920, if we accept the estimate cited; in 1930, if we make some allowance for daughters; and in 1940. The 1920 ratios, 16 and over, are 89.7 percent for females and 89.9 percent for males; the 1930 ratios, 10 and over, 72.2 percent for females and 76.2 percent for males; the 1940 ratios, 14 and over, 82 percent for females, 79 percent for males. For 1940 the estimate based on the assumption is 54 percent of females 14 and over; the Census figure is 57 percent. The separate 1940 data for urban, rural nonfarm, and rural farm areas, shown in the accompanying table, are also helpful in checking the assumption.

PERCENTAGES OF FEMALES, 14 AND OVER, IN THE LABOR FORCE AND IN HOME HOUSEWORK, 1940

Area	In Labor Force	In Home Housework	Total	All Females 14 & over
Urban	31	52	83	100
Rural-nonfarm	21	60	81	100
Rural-farm	12	69	81	100

FEMALES ENGAGED IN UNPAID HOUSEWORK (millions)

	1870	1880	1890	1900	1910	1920	1930	1940
Ages 10 & over	8.9	11.5	14.3	17.3	20.5	23.6	26.4	
Ages 14 & over							25.8	28.9

working age who were gainfully occupied increased, the series rises less rapidly than the total number of females in the labor force. Its movements approximately parallel those in the number of paid domestic servants between 1870 and 1910 (see the table at the end of this paper). From 1910 to 1920 the number of servants dropped, then rose to a point in 1940 only 10 percent above 1910 levels. The number of females engaged in unpaid housework went up rather steadily, reaching a level in 1940 almost 45 percent above 1910.

b) The other main group of persons affected by the definition of gainful work is students. The Censuses of 1850 and 1860 covered students over 15 years of age (even if not also gainfully occupied). All other Censuses omitted them, unless they were also gainfully occupied; the 1940 Census, however, specifically covers student nurses and other students in 'company' training schools receiving some compensation, in money or kind, for attendance. A case for including all persons attending at least professional, business, trade, and technical schools could be made on many sensible definitions of production. This kind of maintenance and expansion of a basic part of our capital is recognized as a real occupation by some pupils, by more parents, and by society at large. If the shift away from the apprenticeship system to the presumably more efficient school, with the resultant cessation of immediate money pay while learning, is ignored, understanding of what has happened to the working population is distorted. For some purposes even students engaged in acquiring a general education may not be omitted; a moment's reflection must show how impossible a modern industrial system would be without literate workers.

Inclusion of students would have enormous effect on the relative importance of the industry 'education', as well as on the aggregate working force in relation to population. In 1940, 9.0 million persons 14 and over were attending school (students al-

ready counted in the labor force because engaged also in some gainful occupation are excluded). Inclusion of students attending schools of higher education alone would approximately double the number of persons 'engaged' in education, as the accompanying figures for 1940 reveal.

Inclusion of all pupils would reduce the growth rate of the industrial category 'education', since pupils per teacher declined between every pair of Census years, and were cut in half from 1870 to 1940 (see Sec. 9). If pupils in schools of higher education alone were

Students (millions)[a]	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940
Attending all schools										
Total	4.1	5.7	6.6	10.0	11.7	13.4	18.0	21.8	27.9	27.3
10 & over[b]					7.9	9.1	12.0	14.0	18.7	19.4
15 & over[b]					2.3	2.6	4.0	4.2	7.2	8.3
Attending colleges, universities, professional & normal schools					0.16	0.24	0.36	0.60	1.10	1.49
Also gainfully occupied										
10 & over							1.1	1.0	1.2	
14 & over									1.0	0.9

[a] Total attending school: 1850–1920, from 1920 Census, *Population*, General Report, p. 1043; 1930, from *Abstract of the 15th Census*, pp. 261–2; 1940, figure reported for age group 5–24 (1940: *Population*, Characteristics, p. 33), stepped up by the 1930 ratio of the total to that age group. The 1850–60 figures do not include slaves, but they were apparently negligible in number; even in 1870 relatively few colored children were attending school. Number, 10 and over and 15 and over: total number, minus students aged 5–9 or 5–14, as given in the 1940 Report, p. 37. Number attending colleges, etc: from *Biennial Survey of Education*. Number also gainfully occupied, 10 and over: Clarence Long; 14 and over, 1930: Long's estimate, minus the Census figure for the 10–13 group; 1940: total attending, 14–24, minus the number in the same age group reported at school under 'employment status'.

[b] The difference between the total and the age group 5–9 or 5–14; it therefore includes a few persons under 5.

included, the growth rate of the total would be raised, after due allowance, of course, for students already covered among gainful workers.

Besides housewives and students, other groups, such as inexperienced workers and inmates of institutions, are of interest in the present connection. These, of far less importance, are considered in the next section.

3 *Also, this definition of gainful work is neither precise nor constant; the figures for some industries are affected more than those for others*

The line between persons engaged in so-called gainful occupations and persons otherwise engaged is rather arbitrary, as we have just seen. But even after accepting the narrower definition of production that is laid down by the Census we run into difficulties. First, the line shifts somewhat from one Census to another (or, in any one Census, from one industrial or occupational area to another). Second, since the position of the line has not been defined clearly, it is subject to vagaries of interpretation by individual enumerators and reporters, and is in consequence really a zone.[12]

a) A good many housewives work also at 'gainful' occupations—almost 4 million were reported in 1930. The general Census rule (the 'priority rule') has been to classify them as gainful workers;[13] no effort is made to divide them between the two classes in terms of some 'full-time equivalent', even if relatively little time was spent at the gainful occupation.[14] Censuses differ in details of treatment. The 1870–90 Censuses instructed enumerators to exclude, from the gainfully occupied, housewives 'without any (or, any other) gainful occupation'. The 1900 Census treated a housewife as gainfully employed if she had a gainful occupation whether she was 'regularly or only occasionally employed'. The 1910 and 1920 Censuses included housewives only if they 'regularly' earned money at their gainful occupation. The 1930 Census also included women so characterized 'unless this (the gainful occupation) takes only a very small fraction of the woman's time'; and, in general, enumerators were instructed when in doubt to exclude from the

[12] These difficulties explain why some workers in international statistics, and sometimes also national statistics, find it convenient to exclude such categories as unpaid family workers, children under 16, and women working on farms, in making space or time comparisons of labor force data.

[13] That is, the intention has been as stated. However, as indicated by the results of the shift in the Monthly Report on the Labor Force schedule, mentioned later, the priority rule has not always been obeyed.

[14] If persons doing unpaid housework were to be included in the domestic service 'industry', then by another Census rule governing the allocation of persons engaged in more than one industry, the domestic service category might be expected to gain at the expense of other industries.

working force persons spending less than the equivalent of one day's work per week on the job. However, while a woman 'who works only occasionally, or only a short time each day . . . shall not be returned as a farm laborer', 'a woman who operates or runs a farm should be reported as a farmer', presumably even if she does not spend much time at such work. In 1940 the criterion of inclusion was *any* work, full-time or part-time, for pay or profit, i.e., work for pay or profit at any time during the week of March 24–30, 1940; yet unpaid family workers (including housewives) on farms were to be excluded if engaged only in 'occasional work'.

The various instructions to enumerators on this problem are not entirely consistent with one another, nor are they always consistent with instructions on other matters.[15] They point a finger at the kinds of work that might have been subjected to variable Census treatment, and that therefore require the attention of those concerned with such work or the industries in which they are significant—seasonal work, unpaid family work, 'gainful' work done at home, and part-time or occasional work done outside the household (the categories are not, of course, mutually exclusive). Violently seasonal industries in which women are of some importance include agriculture, canning, summer hotels, and other types of production with peaks during the summer months; and certain retail stores, with peaks in the spring and late fall. Unpaid female family workers are found largely in agriculture and food stores and eating places. Gainful work at home presumably consists mainly of laundering, dressmaking, and a certain amount of factory home-work. Part-time or occasional outside work occurs largely in domestic service, retail trade, and nursing.

That a large group may be contained in these borderland areas is indicated by the results of the change in the Monthly Report on the Labor Force schedule and instructions in July 1945.[16] The shift from one schedule to another led to a reduction of over a

[15] Such as those for keeping boarders; here, to warrant treatment as gainful work, the occupation had to afford the principal means of support.

[16] See the discussion by L. J. Ducoff and M. J. Hagood in *Labor Force Definition and Measurement*, Social Science Research Council, Bulletin 56, 1947, Ch. II.

million in the number of persons engaged in own-home housework and a corresponding increase in the labor force. Presumably some were, by this shift, allocated to paid domestic service. But most went to other industries; agriculture took a large fraction.[17] If own-home housework were to be treated as a labor force industrial category, however, a fair number would be assigned to it since their other work, although exceeding 14 hours per week, was largely 'incidental' and therefore presumably less important than their housework.

b) A certain number of young people attend school and work as well. As in the case of housewives 'also gainfully occupied', the priority principle has usually led to including them in the working population.[18] But again, as with housewives, some degree of uncertainty arises because of seasonal work, unpaid family labor, and odd jobs generally. The number actually reported is substantial. From the figures cited above it will be noted that the proportion of student-workers shrank between 1910 and 1940.[19] Because the Census dates for 1900 and all except one of the earlier years were June 1, schools were in session fewer days than in recent years, and farming was more important, one would expect a still more substantial percentage of the normal school population to be reported as 'also gainfully occupied'. This expectation would be consistent with the statistics showing a declining trend in the percentage of children in the working force since 1900.[20]

[17] *Monthly Report on the Labor Force*, Sept. 20, 1945.

[18] However, instructions to enumerators of the 1900 Census explicitly require their exclusion if their gainful occupation takes less of their time than their school work (the latter is not defined whether inclusive or exclusive of home or library work). But the 1900 Census was taken as of June 1 when fewer schools were open than there would be today. And as with housewives, the priority rule has probably been violated on occasion.

[19] But this may be due, in part, to the method of estimating student-workers in 1910 and 1920. Nor can the decline between 1930 and 1940 be taken too seriously. The changed treatment of 'seasonal workers' and the elimination of NYA student-workers would cause a decline, probably offset only in part by the changed treatment of 'new workers'. See Durand and Goldfield, 'Estimates of Labor Force, Employment, and Unemployment in the United States, 1940 and 1930', 1940 Census, *Population*, pp. 7–8.

[20] Edwards, p. 92. The 1870 figures, which depart from the trend, are somewhat anomalous; see the later discussion.

The July 1945 change in the Monthly Report on the Labor Force schedule led to a reduction of a quarter million in the number of persons counted as in school and not in the labor force and a corresponding increase in the number counted as in the labor force as well as in school. Presumably the April test preceding it indicated a much bigger shift at that time, since it was held during the regular school year. Apparently here, too, a substantial number of persons may be allocated one way or another, depending on how the schedule is phrased and—it is fair to say—how the schedule and instructions are understood by the reporter and enumerator.

c) Inexperienced workers seeking their first jobs at the time of the 1940 Census were explicitly covered by it. In earlier Censuses, however, there is some question whether new workers were fully covered. It is likely that those with no specific occupation to report may have been disregarded. However, many beginners do have some specific occupation, acquired in school, learned in odd jobs, or picked up from parents. It seems doubtful, therefore, that all new workers were omitted from the 1930 and earlier Censuses. If so, the Durand-Goldfield estimate of 210,000 new workers omitted in 1930 is an overstatement.[21] In any case, the relevant error is small since we are primarily concerned only with *changes* in the number of new workers. Probably they are concentrated in the nonagricultural sphere. There is no reason to believe that the error affects one nonagricultural industry proportionally more than another.

d) Inmates working in penal and mental institutions and homes for the aged, infirm, and needy constitute another group the treatment of which may be and has been variable.

Institutional inmates are mentioned for the first time in the 1900 instructions: they were to be included only if actually engaged in work for which a stated wage in addition to board was received. Beginning with 1940, they were to be excluded in any event. In 1930, it is estimated, gainful worker inmates aged 14 and over numbered about 200,000[22]—less than 0.5 percent of the total gainfully occupied population, though of course a larger but still small percentage of the industrial groups (presumably health serv-

[21] *Op. cit.*, p. 7.

[22] *Ibid.*, pp. 9–10.

ices, agriculture, and government) in which the inmates were classified. In earlier years, probably, worker-inmates were relatively fewer than in 1930, since the institutional population was smaller.

e) Retired and permanently disabled workers were supposed, I gather, to be excluded from all Censuses, though fairly specific instructions appear only in the more recent. Durand and Goldfield estimated, however, that 472,000 were included in 1930.[23] Whether the figure for the years before 1930 would be bigger or smaller than that for 1930 would presumably depend upon how business conditions at the time the Census was taken compared with the spring of 1930.[24] In any case, it is hardly likely that changes in coverage would be large. Part of the resulting error is canceled, as far as the grand total is concerned, by the error arising from the omission of new workers; both errors tend to be magnified with worsening business conditions. Further, the size of both errors may be related to the proportion of employment in nonagriculture, since unemployment of new workers and retired and disabled workers, as well as of other workers, is probably greater off the farm than on it. That is, the error would be smaller in earlier years than in recent.

f) A separate word must be said about unpaid family workers, most of whom are probably also housewives (or children helping in the household) and students, because there is no way of telling how large each group (housewives, students, etc.) bulks among unpaid family workers.

According to the 1940 Census, in which separate figures are shown for the first time, there were about 1.5 million such workers, 1.2 million of whom were in agriculture; the majority of the remainder were in retail trade. The 1910–30 Censuses reported only those in agriculture, namely 1.5 million in 1930 and again in 1920, and 2.6 million in 1910, including seasonal workers, a class not included in 1940.[25] In all these years the reported number of these

[23] *Ibid.*, p. 11.

[24] On the relation between Census dates and business conditions, see Section 7.

[25] Edwards, p. 63. He notes that the 1910–20 figures are underestimates, since they fail to include some unpaid family workers on other than general farms. The persons omitted because of the 1920 undercount are also left out.

workers constituted a very substantial fraction of the total agricultural working force; in 1930 the percentage was 14; in 1910, over 22. Including all seasonal workers (not all were included even in 1930 and earlier years),[26] unpaid family workers may have considerably exceeded the number actually reported—indeed they may have been double that number.

Leaving aside the seasonal question (to be considered later) there is the usual question concerning the exact line that has been drawn between unpaid family workers and persons not counted at all as gainful workers. (This time our question is raised from the viewpoint of a gainful worker category rather than with reference to a category outside the area covered by gainful workers.) Beginning with the 1870 Census, instructions to enumerators mention children assisting in their parents' business. Enumerators were instructed to exclude domestic errands or family chores, and to include only 'appreciable' assistance in mechanical or agricultural industry (retail stores were not mentioned). In 1910 the instructions were revised to exclude, besides general housework and chores, other work at odd times; only 'material' assistance in other than household work was to be covered. In 1930 it is 'regular' work on farms or 'somewhat regular' work in other than farm industry that is to be covered, with at least the equivalent of one day per week in doubtful cases. The 1940 instructions merely require 'actual assistance' on work contributing to the family income. Since the 1940 Census there has been a very significant further change, this time in the instructions to enumerators of the Monthly Labor Force. Beginning with July 1945 a specific, if arbitrary, limit was set on the number of hours spent on incidental chores below which the person performing the work is not to be counted in the labor force. The number of unpaid family workers, especially in agriculture, was thereby increased almost 600,000. According to a pre-test of this new questionnaire in April 1945, the number may be much larger during the regular school year. The increase is, of course, the counterpart of a large fraction of the reduction in own-home houseworkers and students previously noted.

[26] Durand and Goldfield, pp. 8–9.

g) Some corrections have been made by the Census authorities and others for changes in schedules that inadvertently lead to changes in coverage. These, due primarily to difficulties in classifying children and women, suggest the magnitudes that may be involved in some of the problems we have been discussing.

To the reported 1890 Census figure on agricultural workers the Census authorities later added some 600,000 children 10–15. Compared with the labor force propensities for this age group in 1880 and 1900 the 1890 figure seemed low, apparently because instructions on entries for nonworkers were more specific than in the other years;[27] or the other years may have been too high.

The Census authorities felt (1900 Census, p. lxxii, note 1), that "omissions among persons over 15 . . . were inconsiderable and could not be defined more clearly". But on the basis of an analysis of the figures Carson suggests the addition of about 400,000 persons between 16 and 20.

Also because of a change in instructions, and in any case in comparison with 1900, the 1910 figure for children and adult females was felt to be overstated in the agricultural category. About 800,000 persons were therefore deducted by the Census authorities.[28] Clarence Long has tentatively made a further deduction of 650,000 persons, about 250,000 agricultural and 400,000 nonagricultural workers, on the same grounds.

To avoid the kind of overcount that occurred in 1910, the Census instructions for 1920 were modified. The result this time, not only because of the change in instructions but also because of the change in the time of year at which the Census was taken (see

[27] For the first time persons not gainfully occupied were to be reported with respect to activity: housewife, in housework, at school, at home, or with no occupation. Why the added workers were classified in agriculture alone is not clear; but the question is not material, as is indicated in the next note.

[28] Since the 1900 level of agricultural workers was too low because 'laborers (not specified)' were excluded (see below), it is not clear exactly what is involved in the estimate of the 1910 figure. Also, the number of 'laborers (not specified)' that Edwards allocated to agriculture in 1900 depended, in part, on the 1910 level of agricultural workers. The whole business is complicated!

Further, the final 1890 figure for agriculture is actually based on a method Edwards used to interpolate between 1840 and 1910 that was accepted, in part, by Carson. The 1890 correction, therefore, really turns out in the end to be a correction of the aggregate for all industries rather than of agriculture in particular.

Sec. 7), was an undercount. Of the total added by the Bureau of the Census, 820,000 persons, 785,000 were in agriculture and 35,000 in other industries. However, Long questions any undercount in 1920. In his own calculations he uses the unadjusted Census total.[29]

4 *Further, there are inadequacies of enumeration and reporting; these bear more heavily on certain industries than on others*

It remains for us to consider some further questions concerning coverage, namely those arising from inadequacies of enumeration and reporting that do not originate in conceptual difficulties.

a) The first group of persons inadequately or not at all covered in the Census of Occupations (and the Census of Population) consists of workers employed outside the country or with no fixed place of residence—soldiers and sailors afloat or stationed abroad, fishermen, migratory farm laborers or other itinerant workers, some railroad men, etc. Some of these persons are mentioned in a few early Censuses of Occupations, but it is obvious that few are accounted for. Persons changing their residence on Census day and trappers living in the wilds are also in this category.

It is hardly likely that these omissions appreciably affect the over-all aggregates. Certain individual industries or occupations, however, may be substantially influenced, as Daniel Carson points out: transportation, fishing and national defense come to mind.[30]

b) The only undercoverage of total population, and therefore of gainful workers, corrected by the Bureau of the Census is that due to the undercount in certain southern states in 1870.[31] Some 420,000 gainful workers, 3 percent of the revised total, were added and distributed among the various occupational divisions in accordance with the occupational distribution of persons reported in the southern states—that is, mostly agriculture.

[29] See 'The Labor Force in Wartime America', NBER, *Occasional Paper 14*, March 1944, p. 9.

[30] The Census estimates that about 150,000 members of the armed forces were omitted from the 1940 Census because they were stationed outside the continental United States; see Census Release P-44, No. 12, p. 2n.

[31] Edwards, p. 141. We need not concern ourselves with the corrections of the Censuses of Population prior to 1850.

c) In 1940, for the first time,[32] the Bureau of the Census could determine the number of persons for whom no employment status entry had been made, i.e., for whom it was not known, because of carelessness or inability of the enumerator to get information, whether they were or were not in the labor force. It believes that as many as 530,000 persons in the labor force in 1940—1 percent of the total—were omitted because no entry was made for them.[33] No indication is given concerning the industries most affected. On the basis of the 1940 data, the Census Bureau estimates, rather roughly, that some 420,000 persons—or somewhat less than 1 percent—were similarly omitted in 1930.

Omitted entries in earlier years are not known and cannot be estimated. It is hardly likely that they were much less important than those in 1930 and 1940. In the case of 1870, indeed, there is some ground for suspecting that a rather large number of entries for young people were omitted.[34] The percentage of the population counted in the ranks of the gainfully occupied was substantially lower in 1870 than in 1880 and later years up through 1910, as may be seen from the tabulation. The 1870 percentages for males 16 years and over could be expected to be relatively low because of the long death and casualty roll of the Civil War. Conservative estimates are said to put the number of deaths at 600,000. Even after allowance for disabilities and deaths that would have occurred in the absence of war, perhaps 1 percent of all men 16 years and over would be accounted for. In addition, the proportion of adult males in the working force might have been reduced because of the drop in the rate of immigration during the decade preceding 1870. The 'abnormal conditions' left by the Civil War may thus be accepted as at least partially explaining the figures for adult

[32] It would have been possible to determine the number of omitted entries in 1890, because of the requirement that an entry be made for nonworkers; but as far as I know, the number was not tabulated.

[33] Durand and Goldfield, p. 5.

[34] Both Whelpton and Edwards, noting the unusually low percentage of the population reported as gainfully occupied in 1870, use 1880 and 1840 rather than 1870 and 1840 as the bases for estimating the 1850 and 1860 percentages of the population that were gainfully occupied (Edwards, p. 142; Whelpton, p. 342, note r). They ascribe the 1870 situation to the "abnormal conditions following the Civil War".

males. The figures for females seem to be in accord with their trends. But the low percentage of males, 10–15, gainfully occupied, remains unexplained.[35]

PERCENTAGE GAINFULLY OCCUPIED, 1870–1930

	1870	1880	1890	1900	1910	1920	1930
Ages 10–15							
Male	19.3	24.4	25.9	26.1	21.7 (18.0)	16.8 (11.4)	6.4
Female	7.0	9.0	10.0	10.2	8.1 (8.4)	5.8 (5.6)	2.9
Ages 16 & over							
Male	88.7	90.6	90.5 (92.5)	90.5	91.1	91.0 (90.0)	88.0
Female	14.8	16.0	19.0 (19.1)	20.6	24.0 (22.5)	24.2 (24.0)	25.3

Edwards, p. 92. The figures in parentheses for 1890, 1910, and 1920 take into account the corrections suggested by Daniel Carson or Clarence Long.

d) In consequence of the deliberate limitations on the 1820–60 Censuses the figures for some industries are weaker than those for others. Whelpton's attempts to overcome these limitations are discussed in Section 6.

* * * * *

We emerge from the discussion of the aggregates with the keen realization that the gainful worker concept, as used in United States Censuses, is, to begin with, rather hazy; and, further, that the zone of uncertainty surrounding it has been widened by changes in schedules and instructions between successive Censuses. We are left with no very clear notion of its width or its variations from Census to Census. Review of the schedules and instructions merely precipitates qualitative considerations and indicates the *possibility* of variation in count; it offers no basis for quantitative assessment. Except when two counts, utilizing different concepts and procedures, are made for the same period, as in July 1945 (and some earlier pre-tests), and to some extent when

[35] One possible hypothesis concerning 1870 is that this Census date comes closer to a fairly severe trough in business activity than any of the later ones (see Sec. 7). If, as Long suggests, there is a cyclical swing in the percentage gainfully occupied that conforms with business cycles, 1870 would show a smaller labor force propensity than later Censuses; how much smaller is a question, however.

schedules are so arranged that the absence of an entry indicates that something has been missed, as in the 1940 Census, we can merely assume smooth (or, even, little) change, and detect abrupt —and therefore presumably unreasonable—change by comparing contiguous Censuses. The danger here is that our preconceptions will mold the figures with which we end up.

The zones of uncertainty are broader for industrial areas in which are concentrated the groups of the working population subject to variable treatment from one Census to another. Outstanding, of course, is agriculture, though retail trade and personal service also are of concern.

5 *The industrial classification is necessarily gross and rough*

An industrial distribution of the working force derived from Census data is based on information collected not only from individual workers but also from family members, boarding-house keepers, and others responding for them. As a rule little is known about the establishments in which the persons for whom the report is made earn their bread. Consequently, the detailed information needed to distinguish clearly between overlapping industrial categories or industries divided arbitrarily cannot be given. For example, we shall never be able to distinguish clearly, in occupational Censuses, between wholesale and retail trade, since many establishments do both and the reporter can not be sure which is more important; or between manufacturing and trade or service establishments, since it is a quantitative criterion that distinguishes between, say, a small retail bakery making and selling mostly its own products, and a factory establishment selling mostly at wholesale.[36] In the Census of Manufactures, for example, establishments are classified from detailed information on the character and value of individual products.

Even the information respondents might possess is not always fully elicited by enumerators since no detailed industrial or occupational classification is actually utilized in taking the Census.

[36] For similar reasons, the Census of Occupations cannot be as close to an 'establishment' basis as, say, the Census of Manufactures, though it is undoubtedly far closer to an establishment basis than to the enterprise basis on which corporate data are reported in Statistics of Income.

Instructions to enumerators do request detail; Census instructions for 1870 read, "Call no man an 'agent' without further explanation." But it is also true that no enumerator carries around with him a copy of the Bureau of the Budget's standard classifications. Nor, as Carson has pointed out, could he at the wage he receives.

A further difficulty arises from the fact that industrial classifications have changed, with the passage of time, partly because of changes in the economy itself, partly because our ideas on classification have improved. Related is the difficulty caused by changing or indefinite terminology or the changing content of the same and sometimes indefinite categories. The occupation 'clerk' surely has a somewhat different meaning today from what it had in 1870.

As a consequence and at best, Censuses of Occupations can as a rule identify positively only relatively broad industrial categories;[37] and even these must inevitably suffer from fuzzy edges.

Restraining our expectations to a reasonable level, we may inquire how closely Edwards and Carson have been able to approximate the industrial categories from Census data.

a) The initial question concerns the comparability of the 1910 and 1930 industrial categories, the basic framework of the 1870–1930 distribution. These two Censuses were not tabulated by the same code; nor is either classification defined in the detail to which the Bureau of the Budget has accustomed us in recent years. Anyone who has struggled with problems of classification, and especially anyone who has tried to match two sources of data, will appreciate the possibilities of incomparability that lie imbedded, like land mines, even in apparently similar classifications.

As the Census reports give no clue to the comparability of the 1910 and 1930 classifications, Carson cannot settle the issue, al-

[37] This disadvantage should not be minimized. Study of broad industrial groups is of value in getting an initial view and in deriving hints as to forces operating in an economy. But this value is limited. Economists who have examined data for broad, and therefore necessarily heterogeneous, groups sooner or later feel impelled to divide them.

I hardly need say that the conventional (or general purpose) classification, the basis of the Census categories, is really acceptable only when given in detail, and can therefore be utilized in the construction of other, more definitely analytical, classifications.

though he is careful to show how he matched 1910 with 1930 and how his classification compares with the Standard Industrial Classification. Perhaps the failure of the Census authorities to compare 1910 and 1930 indicates serious incomparability.

b) Similar doubts arise in connection with the occupational categories for 1870–1920, the basis for Edwards' and Carson's interpolations and extrapolations. Anyone thumbing through Edwards' detailed notes, or the collation of occupational data for years prior to 1900, given in the 1900 Census, will realize the great variety of categories, the difficulties caused by combining several occupations —and in different ways in successive Censuses—and the frequent vagueness of terminology. Some distinctive occupations simply do not appear in some Censuses, although they existed in the years covered by those Censuses, nor is it clear where they are subsumed.

c) A serious difficulty in building up an industrial distribution of the working population arises from the failure of the Census to obtain the industrial affiliation of a substantial number of workers in years before 1910, or always to publish the information it had. The big groups that therefore straddle more than one industry are 'laborers (not specified)', 'draymen, hackmen and teamsters', and clerical workers. But the problem does not end with laborers, draymen, and clerical workers; even professional persons, cooks, and telephone operators are employed in more than one industry.

Laborers (not specified) were as much as 8 to 11 percent of all gainful workers during 1870–1900. Worse, it is highly uncertain that they may legitimately be assumed to be distributed in some stable proportion or in accordance with any simple formula among all industries. The very vagueness of the occupational category (it is an 'all other' class, including as it does skilled farm hands as well as unskilled workers of various types) militates against such a simple assumption.

The other two occupational categories—draymen, etc. and clerical workers—are more specific, and there would seem to be less danger in distributing them in one way or another among the various industries. Together they are only about half as numerous as laborers (not specified) in 1900 and one-fifth in 1870. All three

occupational groups combined amount to 12–13 percent of the total working force in 1880–1900, and 9 percent in 1870.[38]

There is little one can do in the way of criticizing Edwards' and Carson's distributions of these three (and other) nonspecific occupations. While both tried various methods before deciding on those finally used, neither presents the alternative estimates, nor does either indicate what the industrial picture would have looked like if no effort had been made to distribute the troublesome nonspecific occupations.[39] The unallocated residual in Carson's tables and the clerical occupations in Edwards' tables are not at all estimates of the margins of uncertainty surrounding the industrial distributions. If the residuals included the big nonspecific occupations, or even only portions of them, they would be much larger. By distributing these occupations, both authors have removed some of the uncertainty from the immediate ken of the reader.

As there is little basis for either estimating or commenting on the number of laborers (not specified) allotted to most industries, I shall confine myself to Edwards' method of estimating the number assignable to agriculture. Following Whelpton, he determines the number of agricultural workers in 1870–1900 by interpolating between the 1840 and 1910 ratios of agricultural to total workers[40] with the aid of the ratio of persons living in rural places to all persons.[41] The 1840 ratio of agricultural to total workers is Whelpton's estimate. If it is surrounded by a margin of error (as is suggested, at a later point, that it might be), then so is the esti-

[38] Together with certain other groups treated similarly by Edwards, they amount to 11 percent in 1870, 14–15 percent in 1880–1900.

[39] Edwards does not distribute most clerical workers or a large fraction of draymen. If a percentage distribution of Edwards' data, excluding his category 'clercal workers', is compared with a similar distribution of Carson's, excluding the 'not specified' residual, the discrepancies appear to be rather small (for some purposes!): the broad trends are definitely the same in both.

[40] In 1840 and 1910 there was no large laborers (not specified) group; the agricultural figures are therefore accepted as complete.

[41] The correlation between the ratio of rural to total population and the ratio of agricultural to total working force is good except for the most recent Census, 1940. Better, probably, would be the correlation between the ratio of rural-farm to total population and the ratio of agricultural to total workers; but the rural-farm figures are not available for 1820 and 1840. (The ratio of farm to total families, a good substitute, also is not available for the early decades of the 19th century.)

mate of the agricultural force; and, consequently, that for all other industries as well.[42] It is only fair to note, however, that even if no laborers (not specified) were allocated to agriculture, or even if all were, the downward trend in agriculture's relative importance would still remain very clear and pronounced.[43] For some purposes, on the other hand, for example, if the level or trend of labor productivity in agriculture is in question, it makes a good deal of difference how laborers (not specified) are handled.

d) Carson uses what he calls 'characteristic' occupations as his basic data in estimating the working force of an industry. The characteristic occupations are an excellent basis of estimation for an industry when they have these traits: the great majority of the working force of the industry is in these occupations and the great majority of the people in these occupations is attached to that industry. When they do not have these traits, there is danger that the index of characteristic occupations is biased as an index of the total working force of an industry. If, for example, there has been a tendency, with the passage of time, for the group of professional persons (taken as a whole) to subordinate their positions and accept work as employees of various business concerns and drop strictly professional practice, the group will no longer accurately reflect the trend in the number of professional-grade persons

[42] Since Carson uses Edwards' estimate as a starting point for his own, he also in effect relies on the interpolation method in part.

[43] Similarly, if all women and males under 16 were excluded from the labor force, because of doubt concerning changes in coverage, the downward trend in agriculture would hardly be affected. The figures on the percentage of the working force engaged in agriculture, 1870–1930, tell the story.

PERCENTAGE OF WORKING FORCE ENGAGED IN AGRICULTURE, 1870–1930

	1870	1880	1890	1900	1910	1920	1930
All persons 10 & over							
Edwards	53	49	43	38	31	27	21
Edwards (excl. addition for laborers [not specified])	48	44	39	35	31	27	21
Carson	50	50	42	37	31	27	22
Males 16 & over							
Edwards	57	53	46	40	33	29	25

Edwards' data appear on pp. 98, 104, and 142 of his report.

attached to professional pursuits. As a matter of fact, the ranks of professional-grade persons have been invaded by engineers, chemists, and similarly trained persons frequently—sometimes largely—engaged in industry as employees. Also, with the growth of large scale business, one may expect more and more professionals such as lawyers and architects to be employees of industrial concerns rather than independent practitioners. If these trends have materialized, Carson's estimate of professional service is biased upward, at least until 1910.

Similar questions might be raised in using the statistics for any industrial group. No general bias may be expected, however.[44]

e) I would like to conclude this section with one general criticism. I have mentioned the desirability of using data outside the Censuses of Occupations to aid in assessing the value of the data provided by them. The Census of Manufactures, Statistics of Railways, and other sources should prove useful checks. Indeed, since the Census of Occupations is not adequate in itself, Edwards would have done better if he had utilized other sources in making up his estimates, rather than relying almost wholly on the Census itself. Carson did use a certain amount of other material, but it is not clear that he exploited all possibilities.[45]

I make this criticism with some diffidence. Anyone who has compared various sources will remember the headaches induced by differences in definition, in the way labor turnover affects the figures, and in reporting units (individuals, establishments, or enterprises), not to speak of differences that cannot be identified. The skeptic will profit from a comparison, easily made, of the various estimates of construction employment and labor force by Carson, the National Industrial Conference Board, Kuznets, the National Income Division of the Department of Commerce, and

[44] The problem raised by persons with two or more occupations is also troublesome, since it may raise the figures for some industries and reduce those for others; see Carson's comments. With the growth of large city government, one source of part-time employment has probably shrunk; but the provision of means of rapid transit between rural and urban communities may have stimulated others.

[45] These various sources are not, however, entirely independent. Edwards mentions (p. 33) the use of the Census of Business and of Manufactures, etc. to allocate indefinite returns in the Census of Occupations to the proper industries.

the Bureau of Labor Statistics.[46] We should be grateful for the first steps taken by Edwards and Carson, realizing, as I am sure they do, that the journey is not ended.

6 *For the industrial distributions of 1820–60 the factual foundation is relatively slim*

As has been mentioned, the Censuses of 1820 and 1840 covered only specified industries (in 1830 there was no Census of Occupations); that for 1850 covered only free males over 15; and that for 1860, only free persons over 15. Whelpton estimated the missing persons, a substantial part of the estimated totals.[47] Edwards accepted Whelpton's estimates with very little revision. Since these totals, and their industrial distribution, have been widely used, some remarks on them are not out of place.

Whelpton's method of estimation is rather intricate. Briefly, he first estimated, roughly, the number of slaves and free persons 10–15, in 1860, and these plus free females in 1850, in each industry group except those covered by the Census in 1820. This step gave an estimated total for each industry in 1850 and 1860. Corresponding totals for these industries in 1820 and 1840 were obtained by extrapolating the 1850 and later figures.[48] Using these 1820 and

[46] Or compare Carson's figures on gainful workers in manufacturing and hand trades with Census of Manufactures data on wage earners in manufactures, hand, neighborhood and building trades. Carson's estimate for 1900 is 1,040,000, or about 20 percent, above the Census of Manufactures figure for 1899. His 1870 estimate is 200,000, or 10 percent, above the 1869 figure.

Anyone trying to use the Census of Manufactures for the period prior to 1899 will run into some trouble. The summary data for 1849–99 published in all recent Census of Manufactures volumes suffer from rather annoying defects: they fail to note, unlike the more carefully prepared recent data, significant changes in the scope of the Census and in the definitions of employment.

In the earlier Census of Occupations volumes there is frequent comparison between it and the Census of Manufactures. "The latter suffers by comparison" (*Compendium of the 9th Census*, p. 616); apparently because the number reported in it was smaller than that given in the former.

[47] To 2.5 million persons reported in the 1820 Census Whelpton added 390 thousand, or 14 percent of the estimated total. The corresponding percentages for the other Census years are: 1840, 11; 1850, 31; and 1860, 22.

[48] The 1820 figure for mining was derived from an extrapolation of the 1840 figure.

The extrapolation in each case is of the ratio of persons in the industry to all persons aged 10 and over in 1850 and later years, either along 'a smooth curve' or by simply assuming that the ratio for 1850, or 1850 and 1860, held in 1820 and 1840 too.

1840 figures for industries not covered by the Census as well as for those covered, Whelpton obtained a grand total of the gainfully occupied for each of these two years. He then went back to 1850 and 1860 and estimated grand totals by interpolating on a straight line between the 1840 and 1880 ratios of the gainfully occupied, 10 and over, to total population, 10 and over (see note 34). Using these totals for 1850 and 1860, and the ratio of the agricultural working force to the total (the latter being derived from the relation between the rural and nonrural population), he estimated the agricultural working force. The combined estimate for manufactures and commerce (trade plus transportation) for 1850 and 1860, the difference between the total gainfully occupied and the figures already estimated for the other industries, was split between the two industries on the basis of 1870 and 1840 relations. Finally, 1830 figures were estimated by interpolations between 1820 and 1840.

PERCENTAGE OF POPULATION 10 & OLDER GAINFULLY OCCUPIED, 1820–1930

Whelpton		Edwards	
1820	44.4	1870	44.4
1830	45.5*	1880	47.3
1840	46.6	1890	49.2 (50.0)
1850	46.8*	1900	50.2
1860	47.0*	1910	52.2 (51.3)
		1920	51.3 (50.3)
		1930	49.5

*Interpolated.
Figures in parentheses are Carson's or Long's.
Whelpton's figures are as later revised by Edwards on the basis of adjusted population figures (see Edwards, p. 142); the revision is slight. Since the 1820 and 1840 Censuses failed to specify a lower age limit, it is clear that Whelpton assumes the limit to have been 10, as in recent Censuses.

When Whelpton's figures, expressed as percentages of population gainfully occupied, are compared with later data they do not seem greatly out of line. Since the younger age groups and rural residency were relatively more important in 1820–60 than in 1870 and later years, and labor propensities were smaller in the younger and the rural groups, Whelpton's lower ratios (except in 1870) are at least not in disconformity with expectations.[49] All

[49] Persons in the age group 5–19 accounted for 39.4 percent of the population in 1820, 37.2 in 1840, 37.4 in 1850, 35.8 in 1860, 35.4 in 1870, and 34.3 in 1880 (Thompson and Whelpton, *Population Trends in the United States*, p. 109). Persons living in places with populations of less than 2,500 were 92.8 percent of the population in

this means is, of course, that the figures appear to be more or less what one would expect to obtain by applying, to the sex-age-residency groups of 1820–60, labor force propensities derived from the 1880 or 1880–1900 Censuses.[50]

On the whole, there seems insufficient reason for accepting Whelpton's aggregates for the years before 1870 as anything like precise estimates. Indeed, it is not too much to say that the estimates for missing industries in 1820 and 1840 are almost sheer guesses, and that the straight-line interpolations between the ratios for 1840 (themselves in part, guesses) and 1880 and between 1820 and 1840 are inadequately supported. Compared with the occupation figures as they stood in the published Census reports at the time Whelpton wrote, his estimates are certainly an improvement. The early Census reports are traps for the unwary. But Whelpton clearly states that he considers his estimates merely rough approximations, which he hopes will be superseded by more

---

1820, 89.2 in 1840, 84.7 in 1850, 80.2 in 1860, 74.3 in 1870, and 71.8 in 1880 (Edwards, p. 142).

According to the more recent data, which of course may not be entirely or at all applicable to the situation a hundred years ago, the contemporary figure is low for the rural group because of a low labor force propensity of women—the latter was half of the national average in 1940 (1940 Census, *Population*, II, Characteristics, p. 50).

However, the 1820 figure seems a bit low relative to 1840, perhaps because of an understatement in 1820 relative to 1840. The 1820 schedule asked first for the number of free white persons, second for the number of persons engaged in agriculture, commerce, or manufactures, and third for the number of slaves and free colored persons. Conceivably, the number reported in answer to the second question might cover only free white persons or not cover all colored persons. The 1840 schedule asked first for the number of free white persons, second for the number of free colored persons and slaves, and third for the number of persons employed in mining, etc.—giving less possibility of omitting colored persons.

As for 1840, the schedule called for the number of persons 'employed' in each industry specified. It is conceivable that this might have led to the omission of at least some unemployed persons. But the term 'employed' may really be vague enough to have covered also 'unemployed' persons in 1820.

[50] I say 'more or less' advisedly. If the reported 1850 and 1860 Census figures are stepped up by estimating directly the missing areas, as well as we can, a lower figure is obtained for 1860 (45.7) and a higher figure (48.1) for 1850. Obviously, as the difference between them indicates, the figures cannot be taken very seriously. In calculating them it is necessary to assume, for example, that the labor force propensity of slaves equaled that of colored persons in 1890—a weak reed, at best.

detailed work by the Bureau of the Census. However, his figures for 1820–60 have not yet been improved upon.

Until further work is done, I am inclined to believe that Whelpton's estimates are no better than those obtained by assuming simply that for 1820–60 the ratios of gainful workers to total population 10 years and over lie between 44 and 50 percent. If I *had* to narrow the range, I would put the limits at 46 and 48, and caution the reader against ascribing much validity to them.

If the figures for the missing industries in 1820 and 1840 are untrustworthy, we are left with only the industries for which actual data are available for these years: 'agriculture' and 'manufactures' (1820 and 1840) and 'mining' (1840) seem sufficiently clear and well defined to be comparable with categories given in later years.[51] 'Commerce' (1820) and 'commerce' plus 'navigation' (1840) Whelpton takes to mean what are later called 'trade', 'transportation', and 'finance and real estate'; but the term 'commerce', seems too general for such an identification to be acceptable. The 1840 category 'learned professions' constitutes less than the whole of the later 'personal and professional services', and the figures for them are therefore no more than lower limits.

Slaves, females, and children were rather highly concentrated in certain industries at that time, and any error in Whelpton's allocations would probably be small.[52] However, since manufactures and trade and transportation were estimated by Whelpton as residuals, doubt concerning the grand total carries over to them, and the estimates for these industries might well be expressed in terms

[51] Though there is perhaps a question concerning the degree of coverage of these industries in 1820, especially agriculture; see note 49 above. 'Manufacture' is, in 1820, specifically defined by the Census to include "all those artificers, handicraftsmen and mechanics whose labor is preeminently of the hand, and not upon the field"; i.e., to include the hand trades and construction, and is therefore comparable with the sum of Carson's two categories—'manufactures' and 'construction'.

[52] Of the 2.4 million added by Whelpton for these missing groups in 1850 the largest part—about 2 million—is assigned to agriculture; and of the 2.3 million added in 1860, approximately all. (I assume this after comparing the original Census figures with Whelpton's estimates of the total, allowing for the transfer to agriculture of a large portion of 'laborers [not specified]', originally classified by the Census in nonagriculture.)

of a range. The various figures are brought together in Section 9.

All the figures for the period preceding 1870 suffer from the strictures applying to the later data. Whelpton, too, distributed 'laborers (not specified)' and clerical occupations by rough and ready methods—indeed, much rougher than the methods used by Edwards and Carson.

7 *The data pertain to only a portion of one year in every decade; they tell little about the intervening periods*

Following the fashion, we may conceive of time series as a composite of primary trends, long cycles, business cycles, seasonal cycles, and random perturbations. Obviously it is out of the question for the decennial data we are discussing to yield valid information on anything except trends. But how well can they be expected to do that?

The ease with which the trend of a time series may be seen depends upon the rates of change during each of the three kinds of cycles (i.e., their amplitudes and durations), their regularity, the importance of random perturbations, and the slope of the trend. The steeper the trend, the smaller the composite rate of change during cycles, and the weaker the random perturbations, the more distinct will the trend appear. If, as is the case with the data we are discussing, the entire series is not available, but only occasional observations are at our disposal, our difficulties are multiplied. Even a seasonal fluctuation may obscure a primary trend if the seasonal is relatively sharp, the trend of relatively mild slope, and successive observations are scattered over different months of the year.

a) The seasonal problem arises because not all the Censuses were taken at the same time of year and there is a seasonal movement in the size (and industrial distribution) of the working population.

The monthly count of the labor force, available since the spring of 1940, is distinctly higher during the summer than at other seasons. In 1940 and 1941 the difference between April and the maximum month (July) was about 3.8 million (some 7 percent), most of it concentrated in the age group 14–24.[53] In the 1940

[53] Durand and Goldfield, p. 8. The figures cover persons 14 and over.

Census seasonal laborers out of work in March because of seasonal shutdowns were supposed to be excluded, although some inadvertently were not. In earlier Censuses no specific instructions concerning seasonal workers were issued. Apparently only a fraction reported themselves at the time of the 1930 Census as gainfully occupied, presumably because they were usually doing housework or going to school; Durand and Goldfield estimate that about 1.2 million were counted as gainfully occupied, of whom about 500,000 were student workers. If the estimate for 1930 is accepted, there is of course no problem of passing from 1930 to 1940, as far as the total is concerned.[54]

However, the problem remains for some other years. Census dates were August 7 for 1820, June 1 for 1840–1900, April 15 for 1910, January 1 for 1920, April 1 for 1930, and the week of March 24–30 for 1940. According to the 1941 data in the Monthly Report on the Labor Force, January is about the same as March and lower than April, while both the 1940 and 1941 data indicate that June and August are close to each other and to the peak month, July, and therefore are definitely bigger than January, March, or April.[55] One may therefore expect a discontinuity in the figures between 1900 and 1910 (June 1 to April 15), 1910 and 1920 (April 15 to January 1), and 1920 and 1930 (January 1 to April 1), the effects of which are mitigated because some seasonal workers report themselves gainfully occupied even out of season. In the first two pair of years it would be a decline, in the third, a rise, both mainly in the agricultural working force. Since agriculture was relatively much more important in 1840 than in 1900, even the constancy of the June 1 Census date might not have prevented some seasonal influence on the figures for 1840–1900. Only the 1920 figures have been adjusted by the Bureau of the Census for dis-

[54] Carson has expressed to me considerable doubt concerning the validity of the 1930 estimate; he feels it to be entirely too high.

[55] The 1940 Monthly Report on the Labor Force data are for the week ending June 8—the week of May 11 is much lower and only slightly higher than March or April; and the 1941 data are for the week ending June 14—the week of May 10 is much lower and not much higher than April. It would appear, therefore, that June 1 might not be as much above April 15 as the 1940–41 figures suggest at first sight.

crepancies presumed to have arisen in considerable part from differences in Census dates.[56]

The 1910 figures too were adjusted, but downward, to eliminate what the Census considered the effect of a change in instructions to enumerators.[57] If we are right in thinking that 1910 may have been 'lower' than 1900 because of the seasonal factor, the Census adjustment may really be a net adjustment, the difference between a downward adjustment for the change in instructions and an upward adjustment for the change in dates. Since the Census adjustment applied to agriculture alone, there would seem reason for suspecting that other seasonal industries too, such as building, should be adjusted. Indeed, Long has suggested such a further adjustment.

The inclusion of seasonal workers out of season indicates that the dates of the Censuses cannot be taken literally as the dates to which the responses apply. Some (but apparently not all) seasonal workers not in the habit of working at the time of the Census did report themselves as in the ranks of the gainfully occupied. To that extent, the actual date of reference is uncertain. The number reported is probably greater than the 'correct' figure for the date of the Census and less than the peak figure for the year.

b) The business cycle problem is posed for us because the Censuses have been taken during various cyclical phases as determined by A. F. Burns and Wesley C. Mitchell.[58] Reference dates are on an annual basis for 1840 and 1850, and monthly for the later period. June 1, 1840 came in early mid-contraction, following the peak in 1839; the succeeding trough was in 1843. June 1, 1850 came in mid-expansion, following the trough in 1848, the succeeding peak was in 1853. For the later Census dates the information is more precise: June 1, 1860 came in the middle of Stage IV, the peak being reached in October 1860; June 1, 1870 came at the beginning of Stage VIII, a year after the peak in June 1869, and 6 months before the trough in December 1870; June 1, 1880 came early in Stage III, the preceding trough being in March 1879 and the

[56] Edwards, pp. 138–41.

[57] Edwards, pp. 137–8. The seasonal factor was ignored.

[58] *Measuring Business Cycles* (NBER 1946), Tables 16 and A1.

following peak in March 1882; June 1, 1890 came one month before the peak in July 1890; June 1, 1900 came in Stage VIII, a year after the peak in June 1899 and 6 months before the trough in December 1900; April 15, 1910 came in Stage VI, 3 months after the peak in January 1910; January 1, 1920 came during the peak month of January 1920; April 1, 1930 came 10 months after the peak in June 1929, and 3 years before the succeeding trough in March 1933; March 24–30, 1940 came 22 months after the preceding trough in May 1938 (reference dates for the cycle including 1940 have not yet been set).

On the whole, therefore, the timing of Census dates has varied considerably in relation to business conditions: of 10 dates, 6 came nearer peaks than troughs, 4 nearer troughs than peaks.[59]

The effect of business cycles depends also, of course, on the degree of fluctuation in the number and industrial distribution of the labor force. If the cyclical change is negligible, the fluctuations occurring rather in the ratio of unemployment to employment, there should be little need to worry about the problem—unless one is interested in the number employed rather than in the total labor force. Long's work suggests, indeed, great stability in the total number in the labor force. There must be, however, at least some shuttling back and forth between industries. Consequently, though the total labor force may be more or less stable, its industrial distribution may not be.

Another point is worth mentioning in this connection. The cyclical amplitude of a series is a function of the scope of the series. The more diverse the activities it embraces, the better chance is there of offsetting and reduction in fluctuation. The industrial categories at our disposal differ in the width and heterogeneity of the area covered. We may therefore look for some differences among them in cyclical fluctuation and thus in the degree to which trends may be obscured; also for narrower amplitude in each than in figures for individual industries.

[59] However, according to Frickey's standard pattern of short term fluctuations, 1866–1914, 2 of the latter 4 dates, those nearer troughs than peaks, were near troughs that lay relatively close to the long run average of business activity. See his *Economic Fluctuations in the United States* (Harvard University Press, 1942).

c) Not much is known about long cycles and random perturbations. But we know enough about the former, and about the rapid and far-reaching changes that may be wrought by wars (two Censuses followed large-scale conflicts) to be on our guard. In the case of building construction, for example, 1860 came close to the low in per capita building permits in 1862; 1870, to the following high in 1871; 1880, to the low in 1878. The next peak was 1890 itself, and 1900 the following low; 1910 came close to the peak in 1909, 1920 to the low in 1918; and 1930 came on the downturn between the high in 1925 and the low in 1933.[60] The changes shown by the Census data may therefore at least be questioned as faithful representations of the true trend movements for this industry. Another example: the low level reached by the number of domestic servants in 1920, compared with 1910 and 1930, may reflect the effect of the war.[61] In any case, the trend of this industrial category is obscured.

8 *The figures are for persons in the working force; they provide only approximations to other quantities, such as the number of employed persons*

A gainful worker is either actually gainfully employed or 'actively seeking' work. What 'actively seeking' work means I leave to others. The first question I wish to raise concerns the significance of the industrial attachment of an unemployed person.

a) Many occupations are predominantly associated with a particular industry. Even some apparently rather general or non-specific occupations are really heterogeneous collections of partly or wholly specific occupations. Anyone familiar with bookkeeping and accounting, for example, knows that recording practices vary from one industry to another. A person claiming knowledge, skill, and experience in bank accounting might well hesitate to take a position in the accounting department of a department store. Although the occupational statistics may put all kinds of accountants together, producing the problem of allocation encountered by

[60] Riggleman's data; see Burns and Mitchell, p. 422.

[61] NBER, *Occasional Paper 24*, p. 3. The British figures, cited by Stigler, show a similar low point, in this case in 1921 compared with 1911 and 1931.

Carson and others, an unemployed accountant may feel that he is attached to some particular industry. And in addition to the valuable capital of special training and experience, which is lost when a person moves to another industry, the regional concentration of industry may be an obstacle to movement.

But there is indeed a limit to the strength of the attachment of persons to specific industries. Under the pressure of continued unemployment, geographic obstacles may be overcome, and intangible capital finally written off. Some occupations are only very loosely attached to specific industries: some classes of entrepreneurs and laborers are outstanding in this respect. Multi-occupation persons, working in two or more industries, can perhaps shift their main efforts from one industry to another easily. New workers and immigrants may seize the first opportunity they stumble on. Indeed, there may have been a trend in the strength of attachment to individual industries; but the balance between improved transport, communications, etc. and job simplification on the one hand, and industrial unionization, seniority rules, unemployment compensation systems, immigration restrictions, and other impediments to movement on the other, is difficult to assess.

If unemployed persons have any industrial attachment, the Census figures are relevant to various problems. One is the measurement of the ratio of output to labor input, the latter defined broadly to include unemployed as well as employed workers. An industry so organized that there is much idle time in it may for some purposes be properly charged with the labor not used as well as with that used.[62] In most industries the time of certain classes of labor—clerical, managerial, maintenance—is so treated as a matter of business policy.

b) An important use to which the gainful worker data are frequently put is the measurement of trends in employment and the industrial distribution of employment. The gainful worker data cannot accurately measure either if there is industrial variation in the level and changes in the unemployment rate. At best they are an approximation. How good is it? Unfortunately, little is known

[62] But not necessarily blamed; the basic causes may be outside the control of the entrepreneurs in the industry.

about unemployment in most industries prior to 1930. If we accept certain data in the 1900 Census, we can get some notion of industrial variation in unemployment rates in 1899. These and data from the 1940 Census show that, between the periods covered, the gainful worker data are biased upward as estimates of employment trends and that the bias is greatest (and very substantial) for construction, least for public utilities (Table 1). Admittedly the recent period is unusual, and comparisons of other Census periods might yield quite different results;[63] but too little is known about unemployment in the earlier years to warrant the assumption that it always had negligible influence on the validity of the gainful worker figures as clues to employment.[64]

Another characteristic of the gainful worker data is worth recalling in this connection. All persons are included impartially in the Census of Occupations. Since, by the priority rule, even a part-time or seasonal worker—a student or housewife—is counted as one person, the gainful worker aggregate is larger than a 'full-time equivalent' gainful worker total would be. The relative importance of industries in which there is considerable part-time and seasonal work (for example, agriculture, trade, personal services) will be overstated.[65] And if there is a trend in the proportion of such work, it will affect the relative importance of these industries. On the other hand, as Carson points out, persons with two or more jobs are counted only once in the Census of Occupations.

c) I need hardly warn this audience of the danger in using the gainful worker data as a key to the changing industrial composition of physical output.[66] For narrow industrial classes the danger might be fatal. For the broad groups Carson presents, it is less serious. While there is great variation among individual industries in trends in output per man for categories as wide as agriculture,

[63] Some limited information is provided by the data in Section 7. For a comparison between, say, 1900 and 1930, the bias may well be small for most purposes.

[64] Tabulation of the unemployment data collected in the 1910 Census would add to our information on employment before World War I.

[65] According to the 1939 Census, about a fifth of all workers in retail and service establishments are part-time.

[66] The following remarks apply as well to capital assets, net value added, etc.

TABLE 1

Percentage of Labor Force Employed, by Industry, 1900, 1939, 1940

	Census Year ended May 31, 1900[a]	Calendar Year 1939[b]	Week of March 24–30, 1940[b]
Agriculture	92–95	78–85	93
Forestry & fishing	83–90	61–69	75
Mining	81–90	61–71	82
Manufacturing (incl. hand trades)	91–95	74–81	90
Construction	81–89	48–56	59
Transportation & other public utilities	93–96	83–88	92
Trade, incl. finance	96–98	80–85	92
Domestic & personal service	93–96	71–77	90
Professional service & amusements			
Incl. teachers	87–93	73–82	93
Excl. teachers	96–98	76–83	93
Government	98–99	84–88	93
Not allocated			
Clerical workers	96–98		
Laborers (not specified)	81–89		
Industry not reported		32–45	34
Total	90–95	73–80	87

[a] Manmonths employed as a percentage of available manmonths. Based on the 1900 Census, Tables 2 and 25. The Census authorities consider similar 1890 data inferior in various respects to those of 1900; the 1880 data were not tabulated.

The 1900 data relate to the number of 'months not employed' in the preceding fiscal year. Teachers on vacation were considered unemployed, as would, by definition, seasonal workers not in the labor force out of season. The data were published in frequency distributions with rather wide intervals; for this reason the estimates are presented here in the form of a range, the lower estimate being based on the use of one end of each class, the higher on the other end. Since the data are for monthly units, they may understate the percentage of unemployment in terms of weeks, for it is unlikely that less than half a month of unemployment would cause that month to be reported as one of unemployment.

The National Industrial Conference Board gives average employment in the calendar year 1900 as 94 percent of the total labor force. Our figure, for the fiscal year ended May 31, 1900 (a 12-month period closer to the peak month in general business, June 1899, than the calendar year 1900), is 90–95 percent.

The categories for 1900 are the occupational groups published in the 1900 Census modified to approach Carson's industrial groups more closely.

[b] For the calendar year 1939, full-time manmonths employed (excluding emergency work) as a percentage of available manmonths; for the week of March 24–30, 1940, the number employed (excluding emergency workers) as a percentage of the experienced labor force. Based on the 1940 Census report, *Industrial Characteristics of the Labor Force*, Table 15, and the Census Release, Series P-14, No. 13, Table 2.

The 1939 data are derived from the number of equivalent full-time months worked in 1939 by experienced persons (excluding emergency workers). Emergency work done by them is included. Available manmonths for an industry include the time of emergency workers reporting themselves as normally attached to that industry. Owing to lack of information, it was necessary to assume that the non-emergency work in 1939 of persons with a status of emergency workers at the time of the Census in March 1940 was equal to the emergency work done in 1939 by persons with a status of experienced non-emergency workers at the time of the Census. Since seasonal workers were supposed to be excluded, and the Census was taken in March, there is less overstatement of seasonal unemployment than in 1900. Because the 1939 employment figures are in terms of a full-time equivalent we may expect them to be smaller than the 1900 figures; on the other hand, new workers are excluded in 1939 but not in 1900.

The data are published in the form of frequency distributions by months of work; hence the range. The class for persons not reporting was taken as ranging from zero to 12; excluding these persons would narrow the ranges shown.

The Bureau of Labor Statistics 1939 estimate of average employment (not on a full-time basis) as a percentage of the labor force is 84, excluding emergency workers from the number employed.

manufacturing, etc., the variation is much smaller.[67] Compared with the variation in gainful worker trends, it may in fact be small, making the latter something that may legitimately be called an approximation to relative trends in output. But the degree of approximation is low, and varies from one industrial group to another.[68]

9 *The figures for 1820–1940 summarized; despite their deficiencies they occupy an important place among our historical data*

I now bring together, in Table 2, Whelpton's figures for 1820–70 and Carson's for 1870–1940, with such excisions, additions, and modifications as seem desirable. The reader will, of course, want to consult these writers' papers for their own summary tables, data for industrial subgroups for 1910–40, and various useful derivative tables and notes, as well as Edwards' valuable monograph for its wealth of detailed data and information.

The changes I have made are several. These, together with notes summarizing some of the applicable comments made in preceding sections, are noted below.

First, the changes:

a) Carson points out that his major industry groups were designed to fit, as closely as possible, Kuznets' industrial classification of national income, which itself reflects in part Kuznets' efforts to make the best use of the available data on income payments and business savings. It is for this reason that Carson distinguishes between 'transportation and public utilities' and 'miscellaneous transportation and communication', and places public schools and the postal system in the 'government service' category. I have

[67] Cf. Solomon Fabricant, 'Labor Savings in the United States, 1899–1939', NBER, *Occasional Paper 23*, Nov. 1945.

[68] Indexes in *Occasional Paper 23*, for four major groups show the following. The rank correlation, at least, is perfect!

INDEX, 1939 (1899: 100), RELATIVE TO THE CORRESPONDING INDEX FOR THE TOTAL OF THE FOUR GROUPS

	Agriculture	Mining	Manufacturing	Public utilities
Output	54	123	126	148
Employment	68	114	152	177

TABLE 2

Industrial Distribution of Gainful Workers, United States, 1820–1940

(Unit: 1,000 persons)

	1820	1830	1840	1850	1860	1870 Comparable with: Earlier years	1870 Comparable with: Later years	1880	1890	1900	1910	1920	1930 Comparable with: Earlier years	1930 Comparable with: 1940	1940
Agriculture	2,070	2,770	3,720	4,900	6,210	6,850	6,430	8,610	9,990	10,710	11,340	11,120	10,480	10,180	9,000
Forestry & fisheries	n.a.	n.a.	n.a.	25	50	60	60	95	180	210	250	280	270	120	140
Mining	n.a.	n.a.	15	90	170	180	200	310	480	760	1,050	1,230	1,150	1,160	1,110
Mfg. & independent hand trades	350	n.a.	790	1,260	1,930	2,750	2,250	3,170	4,750	6,340	8,230	10,880	10,990	10,770	11,940
Construction							750	830	1,440	1,660	2,300	2,170	3,030	3,030	3,510
Trans. & other pub. utilities	n.a.	n.a.	n.a.	420	780	1,350	640	860	1,530	2,100	3,190	4,190	4,850	4,810	4,150
Trade							830	1,220	1,990	2,760	3,370	4,060	6,030	6,190	7,180
Finance & real estate											520	800	1,420	1,470	1,550
Educational service							190	330	510	650	900	1,170	1,650	1,630	1,680
Other professional service, & amusements	n.a.	n.a.	n.a.	940	1,310	1,770	140	190	350	500	770	1,080	1,760	1,720	2,320
Domestic service							940	1,080	1,520	1,740	2,150	1,700	2,330	2,550	2,610
Personal service							250	360	640	970	1,520	1,630	2,490	2,500	3,100
Govt. n.e.c.							100	140	190	300	540	920	1,050	1,130	1,690
Unallocated	460	1,160	895	65	80	30	140	195	170	370	600	380	1,340	145*	3,330
Total	2,880	3,930	5,420	7,700	10,530	12,920	12,920	17,390	23,740	29,070	36,730	41,610	48,830	47,400	53,300

Based almost entirely on estimates of Daniel Carson, P. K. Whelpton, and A. M. Edwards as prepared mainly from gainful worker and labor force data collected in the decennial Censuses of Population. See text for detailed explanation and notes.

n.a: not available. n.e.c: not elsewhere classified.

* Difference between number of persons not reporting industrial affiliation (1,335,000), and excess of the 'gainful worker' total over the 'labor force' total (1,190,000).

combined the two transport groups, partly because they are not very clearly distinguishable in the gainful worker data for most years, and partly because the distinction does not seem generally useful. Not needing to put all government activity together, I have taken public schools and the postal system out of 'government service' and have placed the former in a special group of interest in itself, 'educational service', which includes also private schools and other educational activities, and the latter in the 'transportation and other public utilities' category.[69] Because 'trade' and 'finance and real estate' seem difficult to distinguish in the Census data prior to 1910, I have combined them. And because of the importance of and interest in 'domestic service' in connection with the housewife problem, I have broken it out from Carson's 'domestic and personal service' group.[70] The contents of each group are specified in detail by Carson in his Tables 15 and 16.

b) For reasons given above I have identified Whelpton's 'trade and transportation' with the sum of Carson's 'trade', 'finance and real estate', and two transportation groups. Similarly, I have identified Whelpton's 'manufacturing and mechanical pursuits' with Carson's 'manufacturing and hand trades' plus 'construction'; and Whelpton's 'domestic and personal service' and 'professional service' with Carson's two groups bearing similar names plus his 'government service'.

c) I have used Edwards' revisions (p. 142) of Whelpton's 'all occupations' and 'agriculture', which differ but slightly from Whelpton's estimates. The unallocated figure for 1850–70 consists of the difference between Whelpton's and Edwards' totals. For reasons given in Section 6, I have discarded Whelpton's estimates for several groups, 1820–40.

[69] The series for 'educational service' is based on the industrial category as reported in the 1940 Census, extrapolated to 1870 by the number of teachers, including college presidents and professors. The postal system 1910–40 is Carson's series extrapolated to 1870 by means of Edwards' estimates for certain postal occupations.

[70] The domestic service series for 1900–30 is that of Stigler (Table 1) raised 15 percent, as he suggests, and extrapolated to 1870 by the relevant occupational data compiled by Edwards. The 1930–40 figures are Edwards' (p. 84).

d) I have accepted Clarence Long's total for 1910, rather than Carson's, and have adjusted the unallocated figure for the difference between the two totals.
e) Carson's second 1930 total has been adjusted downward. This places the 1930 total on the 'labor force' basis as estimated by Durand and Goldfield. A corresponding adjustment was made in the 'not allocated' item for 1930. Thus modified, the 1930 'not allocated' item represents the net difference between 1,336,000, Carson's figure for the number of persons for which adequate information on industrial affiliation is not given, and 1,191,000, the Durand-Goldfield estimate of the difference between the 1930 'labor force' and 'gainful worker' total.
f) Rounding most of the figures off to the nearest 10,000 accounts for some slight discrepancies between the totals and the sums of the separate items.

And now some notes:
a) The figures for 1820–1930 are for gainful workers 10 years old and over; those for 1930–40, for the labor force 14 and over. (Strictly speaking, of course, the second set of 1930 figures, except the total and the unallocated item, are for gainful workers.) As pointed out above, some of the 1820–70 figures include some minor estimates for young workers, and others are rather indefinite about the precise position of the lower age limit.
b) As mentioned, the grand totals for 1820–60 are Whelpton's estimates or Edwards' minor revisions. If, as I suggest, these estimates are little better than guesses, the reader may wish to replace them with ranges. Assumed labor-force propensities of 44 to 50 percent, for example, would yield totals (in thousands) of:

1820	1830	1840	1850	1860
2,840–3,220	3,800–4,320	5,120–5,810	7,240–8,230	9,870–11,210

There would be correlated changes in other parts of the table, which I have not worked out: in agriculture, 1850–1900; in the sum of manufacturing and construction and the sum of transportation, trade, and finance, 1850–60; and of course in the unallocated item.
c) The figures in the table vary in quality. On the whole, the

columns for 1910–40 are most accurate, those for 1870–1900 less accurate, and those for 1820–60 least accurate. The 1830 data are mere interpolations. Edwards' characterization of his own figures is applicable to Carson's estimates: Edwards states that the figures for the following occupational groups are 'partly estimated': manufacturing and mechanical trades, 1870–1900; transportation and communication, 1870–1900; public service, 1900; professional service, 1870–1900; domestic and personal service, 1870–1900; and clerical occupations, 1900. He calls the following 'largely estimated': trade, 1870–1900; public service, 1870–90; and clerical occupations, 1870–90.[71] He does not qualify the data for 1910 on or his figures for agriculture, forestry and fishing, and mining. The agricultural series for 1870–1920 is, of course, also 'partly estimated', in this case by the interpolation method described in Section 5 and by the adjustments, for under- or over-coverage (Sec. 3), which reflect the difficulties raised by varying the treatment of women and children. To this I would add that Carson's construction group, based as it is on occupations found also to a large extent in manufacturing and hand trades, must be considered 'largely estimated'. The 'other professional service, and amusements' group is also probably in this category for the period prior to 1910.

d) Those who wish to include own-home houseworkers may do so readily; estimates are given in Section 2. Data for students cannot, however, be easily amalgamated with those in the table, mainly because there is no full information on the number of students already counted among gainful workers; see Section 2.

[71] The 1880 figures for clerical occupations are not characterized by Edwards, presumably because of a misprint. I have assumed that he would consider them 'largely estimated', as he does the clerical data for 1870 and 1890.