CHAPTER 3

THE MATERIALS AND THE METHOD

The Meaning of Labor Force and the Materials

ALL the basic statistics on labor force in the United States used in this study were obtained by home interviews, either in complete enumerations of the population by the decennial census, or in sample surveys of representative cross-sections of the population by the census or some other official statistical agency. These statistics are gathered in accordance with a fairly rigorous definition of labor force.

In the United States the labor force is currently defined as the sum of all persons reported by the census to be employed or unemployed during a certain specified week. The "employed" category covers all persons 14 or older who have jobs or businesses for pay or profit, including employers and the self-employed, unpaid family workers in a store or on a farm who help produce a salable product or service, and employees of nonprofit enterprises and government agencies. The "unemployed" category includes persons 14 and older who have no job or business of the above-mentioned sort and are seeking such employment during the survey week.

This definition by the census is not wholly satisfactory from an economic point of view (or indeed from any well thought-out point of view). Among the employed it includes inactive persons: jobholders who are sick, on vacation, weather-bound, or on temporary layoff. The first three of these groups should not be classed as employed or in the economic labor force since they are not producing goods or services and are not currently available for productive effort. The fourth (temporary layoffs) is in the economic labor force because these people are available for productive effort, but they more properly belong under the unemployed heading since they are idle involuntarily and because of economic reasons.[1] Further, persons are included as employed without regard to the number of hours they work. Some put in as much as 90 hours per week, and others, less than 15. Many of the part-time workers want only part-time work and in a sense are only part-time members of the labor force; the others want full-time work and are therefore partly unemployed.

As to the unemployed category the definition not only fails to include the temporary layoffs and partly unemployed workers mentioned

[1] Beginning January 1957, the census reclassified from "employed to "unemployed" persons with jobs but not at work because of temporary layoff or because they were waiting to start new jobs. And persons in the latter category, who happened to be in school while waiting, were classified as outside the labor force.

above, but gives full-time weight to the unemployed who want only part-time work. It is based on no objective description of employability, so that its inclusiveness might vary depending on the quality of the interview, the job market, or the desperation of family need for income. The belief has often been advanced that there may at times be a considerable number of *fringe workers* who want and need work but, discouraged by the difficulty of finding jobs, cease to look and so are not enumerated as unemployed. Conversely, it has frequently been suggested that when family incomes decline many persons may leave school or housework and enter the labor force as *additional workers.*

No facts have been marshalled to suggest that these defects have significantly distorted the size of the labor force; rather, the contrary. Census surveys have indicated that in ordinary times no large number of fringe workers has been overlooked, and the present investigation suggests that the number wrongly left out might be balanced by the number wrongly counted in—persons claiming to be employable or to be seeking work in order to collect unemployment insurance, persons with unrealistic wage and job aspirations, some persons now classed as unemployed on the ground that they would have been seeking a job if well, and other persons on the brink of leaving the labor force.

Studies in this volume indicate that in severe depressions fringe workers discouraged out of the labor force have probably appreciably outnumbered additional workers. Also, in mild recessions since World War II the behavior of the labor force has not been such as to demonstrate the existence of either fringe workers or additional workers. Further studies for scattered months since World War II, when the census gathered special statistics on the preferences of workers for full- and part-time work, have made it possible to construct estimates of what might be called the "economic" labor force—the number of equivalent full-time persons actually at work, plus the number of equivalent full-time persons unemployed for economic reasons (including temporary layoffs). This economic labor force deviated from the reported labor force *seasonally* (because of weather, vacations, and sickness), but over the years the relationship between economic and reported labor force manifested no trend and no systematic response to the recession of 1949–1950 or the Korean conflict. Since such data are available only for scattered months, this volume must rely mostly on the labor force reported by the census; its conclusions apply only to those data and not to a more strictly defined economic labor force. But the results might not be strikingly different even if statistics on labor force were defined and collected according to a more rigorous economic definition.

The concept of the labor force so far discussed refers to recent years.

Some review of the history of the concept and measurement technique is therefore in order (detailed treatment being reserved to Appendixes E and F).

Labor force data have been collected in connection with most decennial censuses since 1820. Through 1860 the data were very inferior in concept, coverage, and accuracy, and in 1870 and 1880, though much improved, were still haphazardly collected and can be used only for rough comparison. Therefore, the figures before 1890 were utilized merely to show that the earlier behavior, so far as can be judged, was not inconsistent with the later results.

The 1890–1930 censuses counted the labor force with reasonable care, but required adjustments for this study, in certain years—to make age groupings uniform in 1890–1910, to compensate for very small undercounts or overcounts in 1890, 1910, and 1920, and to correct for minor differences in the months the censuses were taken (since 1930, always in April). These corrections largely canceled each other. They are not necessarily accurate, and the degree to which they affect the conclusions of the study has been evaluated at relevant points. The chief problem in comparing the "gainfully occupied," as they were called, lay in the fact that the respondent was not asked to specify the exact period of his employment, but was classified as gainfully occupied on the strength of a rather vague "usual worker" status. In April 1940 and July 1945, two innovations in measurement technique were introduced to remedy this vagueness. The first was intended to exclude a number of persons who were usual workers at other seasons of the year but not at the time of the enumeration. The second, made in the census sample survey, was intended to bring in some housewives and students who were defined as belonging to the labor force in the survey week but were being overlooked in practice because their status was different at the moment of enumeration. The census then attempted to adjust earlier data to make them comparable with the new results; but rather intensive analysis in the present study showed that the adjustments were without very solid statistical foundation. The census downward adjustment of the 1930 data to make them comparable with the new 1940 technique is rejected altogether (see Supplementary Appendix H); the census upward adjustment of its 1940–1945 monthly sample data to make them comparable with the technique applied after July 1945 is used because, although no great confidence is placed in the resulting absolute level, most of the month-to-month variations may have been left undisturbed.

In April 1940 the labor force enumerated by the census was slightly smaller than that estimated on the basis of the monthly sample survey (then conducted by the Works Progress Administration) for the same

month. In April 1950 both surveys were conducted by the census, but the labor force as enumerated by the regular census was about 3½ million less than as estimated from the sample survey. Some persons have concluded that the sample was less accurate than the complete enumeration; but a number of tests by the census, including analysis of data from households visited by both surveys, indicate that the discrepancy was attributable to the inferior ability and training of temporary enumerators, compared with the permanent sample survey staff. All of our regular censuses may have thus undercounted, for the relative size of the 1950 labor force does not seem to differ much from that of the earlier censuses. The difference with the sample result is nevertheless disconcerting; and since adjustment is impossible, this study keeps the two sets of data completely separate, using the regular census for analysis of labor force behavior over the long run, and the sample survey for analysis over the short run.

Concerning the usefulness of the regular United States census data for long-run analysis, it may be said that, for all their defects, they have been surprisingly consistent over the decades in both concept and coverage. Statistically, this judgment is supported by many intercensal comparisons during 1890–1950; qualitatively, it is supported by the fact that much the same important groups have been included (or excluded) by all censuses since 1870.[2]

The data for the four foreign countries examined in the present study are probably not as rigorous conceptually, or as comparable over time, as those of the United States. Certainly these nations do far less work of testing and analyzing their concepts and survey methods. No satisfactory data are available in Britain before 1911 (except that data for 1841, 1881, and 1891 could be used for purposes of rough comparison), in Canada before 1911, in New Zealand before 1896, or in Germany before 1895. British data had to be estimated for 1939, as no decennial census was taken between 1931 and 1951. Canadian data have excluded many females in agriculture and school boys working part time, though the latter omissions were offset by the fact that the census was normally taken in June, when many boys were out of school and working on farms. New Zealand censuses are frequently vague on the question of coverage of young children, inexperienced workers, and the disabled. The German census figures suffer from drastic changes in territory as the result of World Wars I and II; and they very likely include in the labor force, as housewives, some who would be excluded from the count in the English-speaking nations. Several of the countries have

[2] The variation in formal coverage has involved minor groups, such as some inexperienced workers not covered before 1940, and some child workers 10–13 not covered in 1940 and 1950.

changed the month in which their censuses were enumerated, and only Canada (in 1951) followed the United States in adopting a more precise time reference in deciding whether a person was in the labor force at the time of the census. On the whole, however, in concept and coverage the censuses of these nations have not changed much over time and have been roughly comparable with United States data during comparable periods—though such a conclusion may not always be accepted by other investigators with different statistical problems.

The usefulness of the United States sample survey of the labor force for short-run analysis depends, of course, on the reliability of the survey methods. It was instituted by the WPA in early 1940 and has been administered by the census since August 1942. It has relied upon a sample of roughly 20,000 households, increased in May 1956 to 35,000, each of which is visited by enumerators for a number of months and is then replaced by a new household—in such a way that not all replacements occur in any one month.[3] The households are not scattered through all the 3,000 or so counties of the nation, but are selected at random from within a smaller number of areas chosen for administrative convenience and as being broadly representative from the standpoint of income levels, occupations, industries, and urban-rural distribution. An effort is made to ensure that the sample reflects any shifts in population. Nevertheless, these monthly estimates of labor force, employment, and unemployment differ from the results that would be obtained by a complete enumeration. The size of the sampling error depends on the size of the sample—the larger the sample the smaller the error. It also depends on the size of the estimate: the sampling error for unemployment will be a smaller number than that for the labor force, but, other things equal, it will be a larger percentage. Finally the sampling error will be greater at some times than at others, depending upon variability of behavior within the sample. Since this degree of internal variability could change every month—and is especially subject to change as the economy moves into recession—the census should, ideally, re-estimate each month its whole schedule of sampling error. However, the census published fresh computations only infrequently before early 1954. The indicated sampling error in the post-World War II period before 1954 was about $\pm 650{,}000$ for a labor force estimate of about 60,000,000—roughly 1 per cent of the labor force. (The percentage error for unemployment, a smaller estimate, would have been larger.) All the errors

[3] These are the households actually interviewed. There have always been several thousand additional households in the gross sample for which interviews could not be obtained because of absence of the householder, vacancy, and other reasons. Currently, about 7,000 households fall into this "visited-but-not-enumerated" category. From January 1954 to May 1956, the number was about 4,000. *Current Population Reports, Labor Force,* Series P-57, No. 168, p. 5.

were larger before 1948. This variability is called the "probable error." The probable error is twice the standard error. The chances are about two out of three that an estimate from the sample would differ from a complete census by less than the standard error. The chances are nineteen out of twenty that an estimate would differ from a complete census by less than the probable error.

Through 1953 there were 68 sampling areas falling within about 120 counties. Late that year the census experimented with a design consisting of the same number of interviewed households (roughly 21,000) scattered through more areas—over 230 in about 400 counties—to reduce sampling error and increase administrative convenience. The old 68-area sample was retained temporarily for purposes of comparison and the results of the two samples, purporting to give more or less independent estimates for the nation, were released for January 1954. The two results were expected to differ by not more than sampling variability. Actually, the January differences were substantially greater than probable sampling error, especially for unemployment. In February 1954 the difference was smaller, though both samples revealed a much greater increase in labor force and unemployment than had ever occurred at that time of year. The census then dropped the old design and continued with only the 230-area sample. Under this sample, the probable error was reduced to, typically, about 600,000 for a civilian labor force estimate of about 66 million.[4]

No change was made in the new design, for the census decided the discrepancy was due, not to defects in the samples, but rather to errors in interviewing. There probably have always been errors in censuses—whether conducted by sampling or by enumeration—arising from misunderstandings or violations of instructions; but during this trial period, errors could have been greatly increased as a result of administering two samples at once with a limited staff. Such difficulties were expected to disappear as the census concentrated its efforts on the new design. There is no way to test the soundness of this judgment; certainly the unhappy experience revealed that, whether because of sampling or of interviewing error, the monthly data may be subject to greater variability than had been anticipated. To reduce this month-to-month variability, though perhaps at the expense of introducing some more persistent bias, the census instituted, in early 1954, what amounts to a new smoothing technique, described as the "new composite estimating procedure." [5] Even so, the errors of month-to-month change have been

[4] *Current Population Reports, Labor Force,* Series P-57, No. 165, p. 7. This is twice the standard error reported in the March 1956 report, shortly before the sample was enlarged to 35,000 interviewed households.

[5] For description of this procedure, as well as a full review of the sampling and survey methods of the census, see *Report of the Special Advisory Committee*

sizable. For example, in April 1956 the error was roughly 360,000 for the civilian labor force and 200,000 for the unemployed.[6] Such errors are very large in relation to the normal change from one month to the next. However, errors are reduced somewhat in this study by bunching the monthly estimates into quarterly averages. The enlargement of the sample by the census in May 1956, to 35,000 enumerated households took in 330 areas and the sampling errors were further reduced—both in degree and in month-to-month change. In August 1956, for instance, the error was 500,000 or 0.7 per cent for the civilian labor force, which was estimated at about 69 million. And the probable error of the month-to-month change was about 360,000.[7]

The methods of constructing the monthly sample survey estimates still leave much to be desired—as is attested by the *Report of the Special Advisory Committee*—but they constitute a notable pioneering effort. They represent some of the best scientific technique yet applied to the problem, planned and administered by one of the ablest groups of experts in the world of statistics.

Only one foreign nation has constructed a sample survey estimate of labor force comparable in quality and period of time covered with that of the United States. This is Canada, where the survey was begun in the form of quarterly estimates in 1945 and has followed American methods closely. Since January 1953, Canada has made a survey one week each month. Canada's sample embraces almost the same number of households as that of the United States although her population is only a tenth of ours. For technical reasons, which need not be described here, sampling errors in Canada are roughly similar in relative size to errors in our estimates. For Great Britain, the short-run analysis of labor force experience during wartime had to rest on annual estimates made in this study from scattered British sources. The wartime analysis for Germany had to rest on estimates by the Strategic Bombing Survey from data which were undoubtedly subject to gaps and duplications.

The Method

Reference was made in the closing pages of Chapter 2 to Paul Douglas' finding of an inverse association between labor force and earnings among 38 large American cities at a given time. Douglas chose large cities as units of study because he had no breakdown by income and he felt that cultural and economic differences among states, counties, or small urban units were so great as to make it impossible to know

on Employment Statistics, Bureau of the Census, mimeographed, August 1954. The members of the committee, appointed by the Director of the Census, were Frederick F. Stephan (Chairman), Lester R. Frankel, and Lazare Teper.

[6] *Current Population Reports, Labor Force*, Series P-57, March 1956, p. 7.

[7] The actual change from July was a decline of 542,000.

whether a difference in labor force participation was truly associated with a difference in earnings.

For example, one small and prosperous town might have a lower labor force participation, not because of its prosperity but because of its remoteness from job opportunities offered in a large city; another town, equally small but unprosperous, might have a large labor force participation, not because people felt forced to work but because of their nearness to a large city where jobs were abundant. Large cities may come closer to providing homogeneity (or other things equal) than the other types of locality for which census data were available.

The *Census of Population* at that time provided no information on earnings of the labor force. Douglas, therefore, had to rely on the *Census of Manufactures* for 1919 and 1929 for an index of a city's earning level. He could not circumvent the difficulty that factory wages might not furnish a satisfactory index of average earnings of the labor force, since the majority of workers are usually in nonmanufacturing industries. An additional complication was that wages were not given separately by age and sex; they were computed by dividing the payroll by the number of workers. Average wages in a city might prove relatively low if its labor force is "loaded" with large numbers of women and children, who almost always earn less than men. Such a situation could create the illusion that many women and children are forced to work because of the low average of wages, when actually the low average wage could be the result, rather than the cause, of a large proportion of women and child workers. Douglas got around this obstacle by dividing the factory payroll by the number of men workers plus a number of women and child earners reduced to an adult-male-earner equivalent. He based the reduction on certain fragmentary information on what women and children earned in relation to men.

When faced with a similar lack of data on earnings of adult males, this investigation followed Douglas' device—for example, in studies of personal disposable income in five nations over time. Douglas also compared earnings both with and without adjustment for inter-area differences in cost of living. This practice is adopted here in inter-area studies; and short-period comparisons over time are made both with and without adjustment for intertemporal difference in the cost of living. But long-period comparisons in this study are made only with income adjusted for the cost of living on the ground that a decade, or a half century, would give workers ample time to see through the "money illusion" and think of their income only in real terms.

Since the cities differed among themselves in size, Douglas compared not the absolute size of the labor force, but the rate of labor force participation. He sought thereby to discover whether a city with

a higher level of earnings, such as Pittsburgh or Detroit, tended to have a systematically different labor force participation rate than one with a lower level of earnings, such as Baltimore. Since persons of different ages and sex have different employment tendencies, Douglas compared first the labor force participation of a particular age-sex group in each city with the earnings of males in the same city; e.g., in Baltimore the labor force of women 25–44, per 1,000 population of women 25–44, was compared with the average wage or salary. A simple correlation coefficient, based on the least squares technique, was then computed for the 38 cities, to determine what the relation was and whether it was significant. Separate correlations with earnings of equivalent adult males were next computed for each of the age-sex groups. Finally, Douglas made a summary comparison between earnings and the labor force participation of the whole population of each city. Before doing this, he had to take account of the facts that each age-sex group has different work tendencies, and each city has a different age-sex composition. Thus a city with a larger proportion of children and a smaller proportion of men may have a lower average labor force participation rate.

In computing an over-all labor force participation rate for a city, these differences in composition are eliminated statistically by standardization. This process, which is the same as fixed-weighting in the construction of price indexes, consists of selecting population age-sex structure of a typical city (or even of the average of all the cities) and weighting the labor force participation rates of the various age-sex groups in each city by that standard composition instead of its own.[8]

[8] Let s_i = labor force of any age group
p_i = population of that age group
$l_i = s_i/p_i$, or the percentage of the population of that age group, which is in the labor force—the "labor force participation rate"
$r_i = p_i/P$, or the standard or fixed ratio of the number of persons, to the number 14 and older
L_m = percentage of the male population in the labor force
L_f = percentage of the female population in the labor force
L = percentage of the population of both sexes in the labor force
$\bar{L}$ = the same percentage standardized

Then $$\bar{L}_m = \frac{\Sigma(l_m \cdot r_m)}{\Sigma r_m}, \quad \bar{L}_f = \frac{\Sigma(l_f \cdot r_f)}{\Sigma r_f}$$

and $$\bar{L} = \frac{\Sigma(l_i \cdot r_i)}{\Sigma r_i} = \Sigma(l_i \cdot r_i) \quad (\text{since } \Sigma r_1 = 1.0).$$

The method is outlined in three steps and illustrated through standardization of the 1940 Baltimore labor force by the composition of Chicago in 1930.

The percentage of each age-sex group in the work force, l_i, is listed below in column one. Each group's percentage is multiplied by its share in the standard population, r_i, in the second column. And the products, l_i times r_i, are added to

So weighted or standardized, the labor force participation rates of the various cities differ from each other only to the extent that labor force tendencies differ and not because population compositions differ.

At least, this is the ideal result. Actually, standardization is subject to two biases. One may be called a *weight bias;* it arises out of the fact that no standard can be completely representative of all the situations to which it is applied and that different standards will in theory yield not only somewhat different average labor force participation rates but also different relationships among the labor force participation rates of any two cities or areas, or of two different dates. The other, called here *interdependence bias,* arises out of the circumstances that the labor force tendencies of a city or area, or at a certain date, may be interdependent with its population composition, and that substituting a standard population composition would yield a spurious or forced result.[9]

---

yield the sums for males, $\Sigma(l_m \cdot r_m)$, females, $\Sigma(l_f \cdot r_f)$, and for both sexes, $\Sigma(l_i \cdot r_i)$, in column three. They have been divided by their shares in the population 14 and older (Σr_i), and the following juxtaposes the standardized and the actual:

Labor Force in Percentages of Population 14 and Older of the Same Sex

	Actual (1)	Standardized ($\bar{L}$)	Effect of Standardization ($\bar{L}$–L)
Males 14 and older	80.5	82.0	1.5
Females 14 and older	32.9	34.8	1.9
Both sexes 14 and older	56.3	58.7	2.4

[9] Standardizing the labor force of a certain locality or date for population composition requires substituting the population composition of some other locality or date, which is called the standard. This process rests on the assumption that the percentage of the population of each subgroup in the labor force would be the same even if it were part of a very different population composition. The assumption need not be true. For example, the number of boys 14–19 has declined greatly during the past half century both in proportion to total population and in ratio to the number of adult males. The tendency of boys to participate in the labor force has also declined enormously. Standardization undertakes to eliminate the first effect—the decline in the relative number of boys in the population—but it is assumed that boys would have the same tendency to work regardless of what proportion of the population they represent.

Suppose, however, that the tendency of boys to work depends on the number of boys in any given family. In 1890, for example, a father with five sons 14–19 may have felt so hard pressed at having such a formidable responsibility that he sent all of them out to seek jobs, thereby encouraging a 100 per cent labor force participation for boys in his family. In 1940, a father of two boys may have felt so affluent because of his relatively light responsibilities for family support that he kept both boys in school, with a consequent participation of zero for boys in his family. Thus to the extent that his feeling of well-being stemmed from having fewer progeny (rather than more income), the decision for or against his children's gainful employment must have been conditioned by the population distribution; in such a case the process of standardization, which couples the 1890 population distribution with the 1940 tendency to be in the

The extent of the weight bias was approximated by experimenting with various standards. These experiments indicated that the weight bias was rather sure to be minor, because the portion of the population differently represented in different cities, areas, or nations—mainly children and old people—did not usually constitute a very large part of the over-all labor force. Since the extent of the interdependence bias could not be ascertained, the next best thing was to present the behavior of both standardized and unstandardized labor force participation rates. The differences in behavior proved small; hence the interdependence bias could not have been large.

Douglas' results for 1920 and 1930 were tested in this study by means of similar correlations among the same 38 cities for one prior, and two subsequent census dates—1900, 1940, and 1950.[10] That is, for those years the labor force participation rates of each of the cities were standardized by the composition of the population of Chicago in 1930. Much the same standardization technique was applied in analyzing labor force behavior among whole states, among the urban and rural areas of states, among different income groups in the same city, among nations, and from one census to the next for the same nation. In the case of the five nations used in the study of behavior over time, the labor force was standardized by the age-sex composition of the United States population in 1940. The United States labor force was also variously standardized over time for changes in the proportion of persons living in rural and urban areas or in the proportion of native whites, negroes, and foreign born. This was necessary since these groups have had very different labor force tendencies at various times and have changed their labor force participation at very different rates over time. Similar additional standardization could not be done for the other countries for lack of data, but the changes in the rural-urban and ethnic composition of their population seem to have been much less than in the United States.

---

labor force, offers a source of error. It is even possible that the composition of the population is in turn influenced by changes in tendency to be in the labor force. For example, wives have always been less prone to work than single women, but in recent years women have been marrying earlier in life and have been having fewer children than the women a half century before. At the same time there has been a rise in the proportion of females who are in the labor force. The process of standardizing for changes in marriage composition rests on the assumption that these two tendencies are independent of one another. But it may well be that a two-way dependency exists: (a) that fewer children mean easier house cares and greater freedom to take jobs, or (b) that the increasing tendency of wives to work, and thereby contribute to family support, in itself explains why couples wed at earlier ages but have fewer children.

[10] This study of the 38 cities omitted 1890 and 1910 because of the difficulty presented by the probability of varying degrees of overcount or undercount in these years in the different cities.

Even in the United States a full cross-standardization for all these changes in population composition was not possible because the rural and urban labor forces were not separately classified by color and nativity in the censuses. However, a device was developed in this study and termed "partial standardization." (See section on *Various Minority Groups* in Chapter 12.) It consisted of "piling" the differences between the labor force participation rate, standardized and unstandardized for rural-urban composition, upon the native white labor force participation rates. Still another device, called "destandardization," was developed to test the stability of participation in the United States and Britain during the nineteenth century, when lack of data on the rural and urban labor forces prevented direct standardization for rural-urban composition. The device (also set forth in Chapter 12) consisted of assuming that the participation rates of the rural and urban populations were the same during the nineteenth century as they have been in recent decades, and then using these assumed rates to compute hypothetical unstandardized labor force participation rates for the whole population at the various nineteenth century census dates. These were then compared with the actual unstandardized rates to see if they yielded similar results.