

Old-Age and Survivors Insurance Records: Derivation of Byproduct Data

by WILLIAM H. CUMMINS*

A VISITOR to the offices of the Bureau of Old-Age and Survivors Insurance at Baltimore, Md., is usually impressed by the speed and accuracy with which his own wage account and his other personal records are located for him by clerks who seem to have no awe of the millions of records and scores of chattering machines through which they wend their way. What the casual visitor may not perceive is that the system that permits such speedy access to the employment and wage records of an individual employee also constitutes an important source of information on group employment patterns and personal characteristics of more than three-fourths of the workers in the United States, Alaska, Hawaii, Puerto Rico, and the Virgin Islands. When these records are tabulated by age, race, sex, industry, and geographic area and by quarterly and annual earnings, the resulting statistical data not only help measure the scope and effectiveness of the Social Security Act but are also of widespread significance to the policy making and economic planning of Government, business, and labor.

As contrasted with the usual methods of collecting statistics, the process by which these data are obtained is so effortless and unobtrusive that it is not unusual for employers to ask: "How on earth do you get such information when we have no record that we have ever filled out a statistical questionnaire for your Bureau, and our tax returns contain no such details?"

Old-age and survivors insurance is just what the title implies. It is a Federal insurance system to which employers, employees, and most self-employed persons—other than farmers—contribute and under which monthly benefits are paid to retired workers, to their eligible dependents, and to survivors of deceased workers.

*Division of Program Analysis, Bureau of Old-Age and Survivors Insurance.

Quarterly reports of wages paid to individual employees and annual reports of earnings by the self-employed are filed with the appropriate collector of internal revenue. He in turn transmits them to the Baltimore office of the Bureau of Old-Age and Survivors Insurance, where the detailed records are kept. Here an individual record is prepared for each of the 90 million account-number holders with earnings in any reporting quarter, and summary records are maintained for the lifetime of each worker. The scope of this record-keeping job is best illustrated by estimates that some 18.5 million tax returns and 220 million earnings items covering taxable earnings of approximately \$135 billion will be received and recorded in 1952.

In 1936, when this Federal insurance system was initiated, it seemed apparent that the records of the Bureau would constitute a "gold mine" of information on wages and employment. Here, for the first time in documented form, would exist a perpetual history of the attachment of workers to certain industries or geographic areas, the migration of others from State to State or from industry to industry, the characteristic ages and annual wages of those employed in selected industries, and other facts of importance in shaping Government and public policies. Important, also, was the fact that these data could be obtained as "by-products" of normal administrative reports, without burdening the employer with questionnaires designed solely for statistical purposes. The advantages of this method of collecting statistics are apparent and are appreciated by both those who must file the returns and those who process them.

Sources of Wage and Employment Data

The three basic sources of old-age and survivors insurance wage and em-

ployment data are shown on the accompanying chart. These consist of Form SS-4, "Employer Application for Identification Number"; Form SS-5, "Employee Application for Account Number"; and Form 941, "Employer's Quarterly Federal Tax Return." In the interest of simplicity, no reference will be made here to any special forms or procedures developed for reporting by the self-employed.

Each employer subject to the Federal Insurance Contributions Act must apply for an identification number to be used on future tax returns. In his application, the employer states both the true name and the trade name of his firm, the nature of the activity, and the location of each place of business. He also gives certain other information relating to the type of ownership (individual, partnership, corporation, or other type of organization) and the origin of the business (purchased as a going concern, for example, or started as a new business). When received in the Baltimore office, these applications are checked to master files to ensure that they are not duplicates. They are then coded for geographic location, industry, type of organization, and origin of business. The appropriate codes are entered in a master punch card that is used in processing future tax returns.

In a similar manner, each employee must apply for an account number, which will be used by his employer in reporting future earnings and under which the Bureau will record his earnings for the remainder of his lifetime. In filling out the application for an account number, the employee enters the date of his birth, his sex and race, and his name and that of his mother and father. Most of these facts are coded and punched into a master employee card for accounting purposes and for future identification and claims determination.

At the end of each calendar quarter, the employer prepares Form 941, the report of taxable wages paid to in-

dividual employees. On this tax return he also indicates the total number of individuals who worked for him during the period (wage items), the total taxable wages paid to them, and

a count of the workers who were employed on the pay period ending nearest the fifteenth day of the last month of the quarter (pay-period employment).

When the return is received in the Baltimore office, individual punch cards are prepared for each employee's wage item. At this point, the employee wage cards contain only the

Basic sources of old-age and survivors insurance wage and employment data

FROM THE QUARTERLY TAX RETURN:

EMPLOYER'S QUARTERLY FEDERAL TAX RETURN

Form 941 (Rev. Jan. 1951)
U. S. TREASURY DEPARTMENT
Internal Revenue Service

1. Federal Income Tax Withheld From Wages (If not required to withhold, write "None") \$ 967 58
2. Adjustment for preceding quarter(s) of calendar year. (Attach explanation. See instructions)
3. Income tax withheld, as adjusted
4. Federal Insurance Contributions Act Taxes (If no taxable wages paid, write "None") \$8140 37
5. Total Taxable Wages Paid (from item 2) \$244 21
6. 3% of wages in item 5 (If employee has and 1/2 of employee to...)
7. Credit or adjustment. (Attach explanation. See instructions)
8. F.I.C.A. Taxes (Item 6 as adjusted by item 7)
9. Total taxes (Item 3 plus item 8). See Depository receipt back on back
10. Total taxes (Item 3 plus item 8). See Depository receipt back on back

11. Name and address of principal place of business and identification number (If incorrectly entered, make any necessary change)
Ralph N. Walker, dba The Guilford Foundry, 10 West Guilford Road, Norwich 12, Conn. 06-9201322

12. Return for Calendar Quarter: March 12

13. Total number of employees employed during quarter: 3
14. Number of employees employed during quarter on 15th day of month: 27
15. Number of business deaths during quarter: 19

16. Has a change of ownership or other transfer of the business taken place during the quarter? Yes [] No []

17. If "Yes," attach statement showing name and address of new owner or processor.

18. If "Yes," attach statement showing name and address of new owner or processor.

EMPLOYEE'S ACCOUNT NUMBER (If number is unknown, file with Office of the Collector & or A)	NAME OF EMPLOYEE (Please type or print)	Taxable WAGES Paid to Employee During Quarter (Before deduction)	State Pension or Unemployment Compensation (Check item 8 "I")
(1)	(2)	(3)	(4)
991 29 7648	John Norman Thompson	285 67	Conn.
020 41 2342	Wayne A. Smith	213 70	"
727 22 9864	Earl K. Farmer	360 50	"
303 36 2754	Charles Graynor	247 90	"
580 29 0430	William H. Shaw	139 87	"
001 19 1053	Samuel Arey	243 50	"
665 39 1108	Thomas L. Terry	120 45	"
002 01 0011	Gordon Ross	476 60	"
115 10 2053	John White	355 75	"
234 65 7654	Raymond White	342 20	"
009 98 8765	Peter M. Claudy	112 55	"
232 22 3458	Robert Monroe	337 50	"
004 72 3690	Harry Korelli	75 00	"
001 35 0005	Ruth K. Barish	135 00	"
Total		3446 19	

21. TOTAL FOR THIS RETURN—Total taxable wages paid during quarter \$ 8140 37
22. Total taxes (Item 3 plus item 8) \$ 244 21

THIS FORM MUST BE FILED WITH U. S. COLLECTOR OF INTERNAL REVENUE ON OR BEFORE THE LAST DAY OF THE FIRST MONTH FOLLOWING THE CLOSE OF THE QUARTER EVEN THOUGH NO TAX IS PAYABLE.

1. Employer Identification
2. Employee Identification
3. Individual Earnings for the Quarter
4. Total Taxable Wages Paid
5. Total Number of Employees
6. 15th Day, Last-month-of-quarter Employment
7. Business Deaths

FROM THE EMPLOYEE APPLICATION:

APPLICATION FOR SOCIAL SECURITY ACCOUNT NUMBER
SPECIMEN
991 - 29 - 7648

1. EMPLOYEE'S NAME (Last, first, middle initial) John Norman Thompson
2. ADDRESS (No and P.O. Box) (City, State) 1022 Carver St., Boston 56 Mass.
3. POINT OF BIRTH (City, State) John Norman Thompson
4. AGE IN MONTHS (Date of Birth) (Date) 30 April 15, 1907
5. RACE (Specify) Hadley Hampshire Mass.
6. SEX (Specify) Male
7. FEDERAL INSURANCE CONTRIBUTIONS ACT (Check one) (Specify) John Willis Thompson
8. TYPE OF ORGANIZATION (Specify) Unemployed
9. BUSINESS DEATHS (Specify) None
10. DATE (Month, Day, Year) Aug 4, 1937
11. SIGNATURE (Typed name) John Norman Thompson
12. RETURN COMPLETED APPLICATION TO NEAREST SOCIAL SECURITY ADMINISTRATION FIELD OFFICE

1. Employee Name and Account Number
2. Age
3. Race
4. Sex

FROM THE EMPLOYER APPLICATION:

EMPLOYER APPLICATION FOR IDENTIFICATION NUMBER
SPECIMEN
06-9201322

1. EMPLOYER'S NAME (Last, first, middle initial) Ralph N. Walker
2. ADDRESS (No and P.O. Box) (City, State) The Guilford Foundry, 10 West Guilford Road, Norwich 12, New London Conn.
3. TYPE OF ORGANIZATION (Specify) Foundry - Mfg. Gray Iron
4. BUSINESS DEATHS (Specify) 20
5. DATE (Month, Day, Year) Oct. 10, 1936
6. NAME OF BUSINESS (Specify) Foundry - Mfg. Gray Iron
7. TYPE OF BUSINESS (Specify) Castings for sale and on job-order basis
8. SIGNATURE (Typed name) Ralph N. Walker
9. DATE (Month, Day, Year) 7/1/51

1. Employer Name and Identification Number
2. Geographic Location
3. Industry
4. Type of Organization
5. Business Births

employee's name, account number, and taxable wages for the quarter. An employer total card is likewise prepared, in which is punched the name and identification number of the employer, the calendar quarter, the number of wage items appearing on the return, the total taxable wages, and the pay-period employment. The individual wage cards are then summarized, and the total number of such cards and the total wages are compared with the corresponding figures on the employer total card. If the amounts are identical, the tax return is said to be "in balance," and the employer and employee cards are released for further processing.

When these punching and balancing operations have been completed for all tax returns received for a calendar quarter, two separate files of punch cards have been created. The first of these, the employer total cards, are mechanically matched (or collated) to the employer master-card files that contain the industrial and geographic codes for each employer. In this operation, the codes are transferred from the master card to the matching total card. The total cards now contain the name and identification number of the employer, his industrial and geographic codes, the total wages, total wage items, pay-period employment, and the calendar quarter covered by the report.

The employee wage cards are accumulated for four consecutive quarters, sorted in account-number sequence, and then mechanically matched with the employee master-card file that contains the age, race, and sex codes for each employee. A sample¹ of employee wage cards is then selected for statistical purposes, and a new deck of statistical cards is prepared that omits information not essential to later statistical operations. These new cards, which might be called "worker cards," are then sorted by employer identification number and collated to the employer master-card file from which the appropriate industrial and geographic codes are obtained. Each worker card in the sample now contains the account number of the employee; his

age, race, and sex; his quarterly and annual taxable wages; the identification number of the employer; and the industrial and geographic codes.

Types of Available Data

It will be apparent that there are two broad categories of wage and employment data (as distinguished from claims and benefit statistics) that are available from these two sets of records.

The first of these, "employer statistics," provides information on the number of employing establishments, their births and mortality, and their wage payments and employment experiences, distributed by industry and geographic area. As a measure of employee distribution and business population, these data are valuable to business firms in determining areas of potential sales and quotas of established sales offices, and in conducting similar market research; to Government they have been useful in estimating the salary and wage component of national income, as a control on the accuracy and representativeness of statistical studies made on a sample basis by the Bureau of the Census and other agencies, and in defense planning. This type of information has widespread significance in economic analysis, and general public uses for the information have grown rapidly in recent years.

The second category, "employee statistics," is broad in scope and of diverse utility. At present, employee statistics find their greatest usefulness in analyses of current and proposed provisions of the Social Security Act; in studies of quarterly and annual earnings, guaranteed annual wages, and allied problems of employment duration and income maintenance; in research on problems of mortality and health, when correlated with data from other sources; in employment and manpower studies; and in formulating industrial pension plans. Eventually these data may have even greater value in general economic planning, for they provide information on employee earnings, by age, race, sex, industry, and geographic area; the recency and continuity of employment; mobility of workers in terms of geographic areas, industries, or individual employers;

annual earnings by industry and area; and similar important characteristics of the labor force.

Many years have been expended in developing these data to their present level of usefulness, and many more will be required to exploit their full potentialities. Some of the more difficult problems encountered in developing the industrial and geographic breakdowns of these data, and the unique procedures devised to meet them, are described in the following paragraphs.

Industrial and Geographic Classifications

It was recognized from the beginning of the program that the Bureau's wage and employment data would be of limited value unless they were classified by industrial activity and geographic location. Since geographic areas were so well defined, the selection of an appropriate code was dependent on the amount of detail ultimately needed in the tabulations. In the first few years of the program the Bureau used a geographic code consisting of six numerical digits, which identified the States, counties, and individual cities and towns. This code was supplanted by a four-digit classification, which separately identifies States, most counties, and cities that have county status.

The development of an appropriate industrial classification was a task of far greater magnitude, since no comprehensive and detailed code existed at the time. Numerous classifications had been developed by Government agencies for the industrial areas with which they were concerned, such as the Bureau of the Census classification of manufacturing establishments and the Bureau of Mines codes for mineral industries. Most of these codes, however, contained overlapping, partial, or obsolete groupings or large miscellaneous categories that reflected lack of interest in fringe industries or primary concern with the maintenance of established historical series.

As a first step, therefore, it was necessary to construct a complete industrial classification that would be applicable to the wide range of activities covered by the Social Security Act. To achieve comparability with

¹For a description of sampling methods see the *Bulletin*, June 1951, pages 17-19.

the statistics produced by other agencies, it was desirable to utilize existing codes wherever feasible and to resolve areas of conflict between such codes. Early efforts initiated by the Social Security Board were soon joined with those of the Division of Statistical Standards, Bureau of the Budget, which then sponsored the development of a standard industrial classification.

In the interim, from 1937 until 1941, the Bureau of Old-Age and Survivors Insurance used the Social Security Board classification, which at that time consisted of about 70 major industry groups. While such codes were applicable to the majority of employer reports, the resulting data fell short of meeting administrative and analytical needs. Consequently, in 1942 questionnaires were sent to all covered employers to obtain current information on the nature of their business and the addresses of their establishments. At that time, the approved Standard Industrial Classification for manufacturing industries and the tentative code for nonmanufacturing industries promulgated by the Bureau of the Budget were applied to the newly refilled records. First steps toward a program of interagency coordination were taken when these coded questionnaires were transmitted to the State employment security agencies for incorporation of the codes in their records.

Establishment Reporting

About the same time, another important step was taken to correct one of the most serious obstacles to the production of accurate wage and employment data. Treasury Department regulations require that each employer file a consolidated report covering the employment of the entire firm. This method was satisfactory in handling reports filed by nearly 2 million "single-unit" firms that operated only one place of employment and whose total employment represented about 60 percent of that reported by all firms in a given quarter. It created a difficult statistical problem, however, with respect to some 20,000 "multi-unit" firms that operated nearly 200,000 establishments and whose quarterly employment accounted for the remaining 40 per-

cent. Some method was needed to obtain a breakdown of the multi-unit employment and wages that were allocable to various industries and geographic areas.

In 1941, the Bureau of Old-Age and Survivors Insurance began making personal contacts with these multi-unit employers to enlist their cooperation in the adoption and use of a voluntary procedure called "establishment reporting." This plan involved the grouping of employee names by establishment within the consolidated tax return, the identification of each group by means of an employer-assigned number, and the filing of a master list of establishments with the Bureau. In effect, the plan provided a separate tax return for each such establishment, thus permitting more detail in the Bureau's industrial and geographic data. Contacts with the employers were highly favorable; more than 95 percent of the firms agreed to cooperate in the establishment reporting plan, and many who had originally declined to participate have since volunteered to do so.

Establishment reporting should not be construed as a statistical technique that is free from error. In most instances, individual payrolls permit accurate listings of employees for each of the establishments operated by the employer. This is particularly the case in the mining and manufacturing industries, or where the firm's various places of employment are widely dispersed geographically, large in size, or functionally dissimilar. The technique is not always applicable to the reports of some employers who maintain "confidential" payrolls on which are carried the names of executives who may actually work at various establishments of the company. A problem also exists in some of the retail and service industries, where numerous small establishments and a high incidence of employee transfers may make it impossible for the employer to show employee names by individual establishments. Such deviations from strict "establishment" reporting are relatively insignificant, affecting primarily the employer data on number and size of employing establishments and creating some minor inaccuracies at the county level in geographic tabulations of employee data.

Rotational Refiling

Once the employer had received his identification number, he had no obligation to notify the Bureau when changes occurred in the nature of his business or in the geographic location of his establishment. Some periodic recontact with him was therefore necessary to maintain the accuracy and currency of industrial and geographic codes. Questionnaires sent to the whole body of employers at long-time intervals did not seem to be the answer, for such projects were costly and voluminous in scope and created peak loads for the Bureau that might interfere with its regular administrative functions. Moreover, while periodic mass refiling would result in high accuracy of classifications at a point in time, a steady but unknown amount of deterioration in codes would take place during the intervening years.

Comparative studies indicated that there were varying rates of change in the type of industrial activity and the physical location of establishments. Mining establishments, for example, by their very nature showed little or no change over a period of years, and to a lesser extent the same was true of the heavy manufacturing industries. In retail and wholesale trade, however, and in the personal-service industries such changes were more frequent and might have an important effect on the accuracy of the industrial statistics.

From these studies a system of "rotational refiling" was evolved, under which questionnaires were sent to selected industries at intervals of from 1 to 5 years, depending on the frequency of the changes occurring in the industry. To obtain closer coordination with the statistical work of the State employment security agencies and to avoid duplication of contacts with employers, arrangements were made to supply the agencies with copies of employer questionnaires under procedures that permitted the States to challenge industry codes assigned by the Bureau.

Pay-Period Employment

Until recently, one additional problem remained unsolved with respect to the quarterly employer data. Since 1937, the employer's tax return

has contained a question relating to pay-period employment. The query currently appears on Form 941 as item 14 and asks for the "number of persons employed during pay period ending nearest the 15th of third month in quarter." This is the only statistical question asked of employers each quarter on Form 941, and it is an attempt to obtain a count of employment as of a given point in time as contrasted with the total number of employees during the 3-month period. The question is a standard one, approved by the Bureau of the Budget for use by State employment security agencies, the Bureau of Labor Statistics, the Bureau of the Census, and other agencies.

Since the answer to item 14 represented the pay-period employment for the firm as a whole, data for different geographic areas and industries could be determined with accuracy only for the single-unit firms. In the absence of pay-period figures for the individual establishments of multi-unit firms, the total pay-period employment was distributed by the Bureau to each reporting unit shown on the tax return in the same proportion as the wage items listed thereunder. Because the Bureau was unwilling to increase the burden of statistical reporting on employers,

this technique was followed for many years, with the full realization that it was not reliable since it did not allow for variations in labor turn-over, or for seasonality, catastrophe, or similar factors of unequal effect on the pay-period employment of different reporting units.

The growing importance of the data brought increasing pressures to obtain more accurate employer-reported figures. Consequently, early in 1952 personal interviews were held with a small sample of multi-unit employers to determine the additional effort involved in supplying an establishment breakdown of the total pay-period employment figure, and whether the firms would be willing to itemize these figures on the recapitulation attached to their quarterly tax returns. The response was uniformly favorable; each firm advised the Bureau that the answer to item 14 was a summation of individual figures collected from their various establishments, and the only additional effort would be that of copying these figures from work sheets to the quarterly report form.

With this encouragement, letters were sent to 17,000 multi-unit firms that were using establishment reporting methods as of March 1, 1952. The letter described the problem and

the solution proposed by the Bureau and asked for the reaction of the employer and his cooperation if feasible. Although the Bureau has always enjoyed friendly and cooperative relationships with employers, the scores of favorable and complimentary letters that poured into the Baltimore office were perhaps without precedent.

The psychological stimulus that resulted in such unexpected reaction is not one that lends itself to statistical evaluation alone. Within 2 months, about 9,000 replies had been received from these firms, more than 97 percent of whom agreed to furnish pay-period employment by establishment. While it is hoped that this attitude will also prevail among the employers who did not reply, the complete story will not be known until the tax returns for the first quarter of 1952 are received in Baltimore. An individual check will then be made to identify those cases in which no answer was received and no actual cooperation shown.

The problems that remain to be solved are primarily those relating to statistics produced by other Government agencies. Efforts to achieve greater comparability and a maximum of coordination with these data will be discussed in a later article.

Notes and Brief Reports

Social Welfare Expenditures, United States and Great Britain, 1949-50

All governments profess to seek the improvement of the economic and social well-being of their people, but agreement on measures of progress in achieving this goal is far from universal. Welfare values differ, as do the programs through which they are expressed. The same terms mean different things in different countries.

Largely through the efforts of international agencies, progress is being made toward the development of comparable indexes of the general well-being of the population. Such measures as expectation of life at birth, literacy, percent of unemployment, and per capita income have come to

possess a common meaning in most countries. At the same time, important advances have taken place, country by country, in the collection of the basic data underlying these indexes.

Another type of international comparison in this field involves the measurement of government effort as expressed, for instance, in expenditures for social welfare as a percent of national income or as a percent of government expenditures for all purposes. These ratios are useful gauges of national concern with social problems, although it should be recognized that the results are affected by such factors, among others, as the age composition of the population, the division of responsibility between public programs and voluntary effort, the coverage and relative maturity of

the income-maintenance programs and the acuteness of housing and other problems left as a legacy of the war.

The present note compares public social welfare expenditures in the United States and Great Britain in relation to national income and government expenditures for all purposes. The data relate to the fiscal year 1949-50, the most recent year for which fairly complete information is available, and include, in the United States, expenditures by Federal, State, and local governments, and in Great Britain, expenditures by the National Government and by local authorities. Social welfare, as used here, is broadly defined to include the income-maintenance programs (social insurance and related programs, public assistance, and family allowances), health and medical care programs, education, such welfare services as vocational rehabilitation and child care,