

---

# Commentary: Continuous Work History Sample

By Warren Buckler\*

---

The Continuous Work History Sample (CWHS) is the oldest major continuous longitudinal sample data source in the Federal statistical system. It was developed to fulfill the need for statistics to be used in planning and operating the old-age insurance system established by the Social Security Act of 1935. This Act led to the establishment of one of the most extensive sets of administrative records in the world.

Before the 1930's, there was very little national information on earnings and employment. What existed was collected by various government agencies, cross-sectional in nature, and limited to a particular time period with no ability to trace an individual's earnings over a long time span. The need for statistics on the continuity of earnings and employment data became urgent during the 1930's when widespread unemployment became a serious economic problem with significant policy implications. Several surveys were made using interviews to obtain work histories. For the most part these surveys were limited in coverage and were based on the memory of individuals being interviewed, which cannot be relied on for accurate wage and employment information over long periods of time.

With the enactment of the

---

\* Director, Division of Statistical Operations and Services, Office of Research and Statistics, Office of Policy, Social Security Administration.

Social Security Act, the development of comprehensive and accurate data on the continuity of earnings and employment became possible. The quarterly reports of taxable earnings of individual workers that employers were required to file became the base for a system of wage recordkeeping from which the total taxable wages received by an individual from various employers could be determined and patterns of earnings and employment of a worker over an extended time could be traced. It was recognized immediately that these records had value for broad economic and demographic research purposes as well as for administering the old-age insurance system, which was the responsibility of the Bureau of Old-Age and Survivors Insurance (BOASI), the predecessor of the Social Security Administration (SSA).

In 1938, BOASI began tabulating wage record data on 1937 annual earnings by age, sex, race, and the State in which the worker was employed. These annual tabulations were repeated for 1938 and 1939. Additionally, the data for 1937-39 were tabulated by industry and cumulative wages. The 1937 and 1938 data were tabulated on a 100-percent basis; data for 1939 were based on a 20-percent sample. When, in 1939, the Social Security Act was amended to require continuity of employment and wages as a basis for determining benefit

eligibility and amounts, the type of statistics required by BOASI was affected: The number and specified quarters of coverage were now needed along with the cumulative wages.

In 1940, a small (0.1 percent) sample of Social Security records of individuals in 22 States was selected. The following year, the first CWHS, designed along the lines that are known today, was established. The original sample was intended to include 4 percent of all living wage workers whose accounts were established during 1937-40. Selection was made on the basis of certain digits in the Social Security Number. Each year, new workers were to be added based on the same digits of the Social Security Number. The current year's wage record for each worker in the sample was to be added to the accumulated information, and the accounts of deceased workers and those who became entitled to primary insurance benefits were to be segregated from the others. Because of the wartime need for economy in the use of personnel time, the sample was limited to one-fourth of the originally planned size (1 percent).

The **Social Security Bulletin** reported on the progress of and findings from the CWHS. A 1951 article by Jacob Perlman, reprinted in this issue, presents summary statistics and an analysis of the data from the 1937-48 CWHS, a 12-year accumulation of data under that

Continued on page 56

**Table M-29.—AFDC, Emergency Assistance: Number of families, by State, July–September 1986**

State	July 1986	August 1986	September 1986	State	July 1986	August 1986	September 1986
Total .....	34,726	41,978	43,900	New Jersey .....	590	573	602
Arkansas .....	571	0	0	New York .....	3,032	12,527	7,246
California .....	669	662	683	Ohio .....	5,515	6,589	7,183
Delaware .....	220	196	244	Oklahoma .....	703	1,093	1,489
District of Columbia .....	964	1,086	1,734	Oregon .....	1,716	1,716	1,754
Georgia .....	821	420	442	Pennsylvania .....	45	59	44
Illinois .....	265	138	191	Puerto Rico .....	339	311	1,653
Kansas .....	138	148	208	Vermont .....	314	293	330
Maine .....	326	399	257	Virgin Islands .....	3	3	2
Maryland .....	2,392	2,201	1,855	Virginia .....	20	12	11
Massachusetts .....	5,324	5,168	6,448	Washington .....	645	655	682
Michigan .....	5,364	3,940	5,903	West Virginia .....	2,359	735	2,494
Minnesota .....	1,602	1,505	1,388	Wisconsin .....	19	688	202
Montana .....	92	92	118	Wyoming .....	511	518	463
Nebraska .....	167	251	274				

CONTACT: Emmett Dye (202) 245-2743 for further information.

## Commentary

Continued from page 12

system. A full-scale description of the CWHS, co-authored by Perlman and Benjamin Mandel, was published in the February 1944 **Bulletin**.

Through the years, the CWHS system has been adapted to keep pace with the numerous changes in the Social Security program. As the Social Security Act was amended, the CWHS processing, files, and record contents were modified to reflect the changes. A subset of the 1-percent sample (a 0.1 percent sample) was developed to meet the needs for more extensive earnings detail without overburdening the 1-percent sample file. In response to increasing interest in obtaining data on employee-employer characteristics, annual files of the 1-percent sample workers with wage employment in the year were created. These new files contained data on the employee's age, sex, race, wages paid by the employer, and corresponding employer characteristics (industrial activity and

geographic location). These annual employee-employer records have been retained as part of the system since 1957. As interest increased in the area of measuring economic, industrial, and geographic mobility, so did the demand for a longitudinal file that would contain such data. In 1969, the records from the separate employee-employer files for 1957-66 were merged to permit all of the records associated with a particular employee over the entire time span of the file to appear together. This longitudinal employee-employer data file is updated annually with the current year's data.

The CWHS is a major source of Social Security program statistics and workforce data. It is a rich resource whose great potential beyond its use in SSA has been recognized. From the early 1960's through the mid-1970's, the CWHS became an increasingly popular resource among researchers outside SSA,

and special versions of the file were made available to them under the constraints of strict confidentiality and agreements on the specific conditions of their use. Annual users conferences were held to discuss the ways in which the file had been used in research and to share information about the data. However, because of certain confidentiality and disclosure issues raised in the interpretation of the Tax Reform Act of 1976, release of the CWHS to non-SSA users has been discontinued.

Today, the CWHS is used extensively agencywide to provide descriptive program statistics, to support our research and analytical responsibilities, to analyze the impact of recently enacted legislation, to evaluate proposed new legislation, and to provide data used in actuarial estimates. Since its inception, the CWHS has been an extremely important data source and it will continue to play a vital role at SSA.