

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: An Appraisal of the 1950 Census Income Data

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

Volume Publisher: Princeton University Press

Volume ISBN: 0-691-04102-4

Volume URL: <http://www.nber.org/books/unkn58-2>

Publication Date: 1958

Chapter Title: Coordination of Old-Age and Survivors Insurance Wage Records and the Post-Enumeration Survey

Chapter Author: B. J. Mandel, Irwin Wolkstein, Marie M. Delaney

Chapter URL: <http://www.nber.org/chapters/c1052>

Chapter pages in book: (p. 167 - 176)

# Coordination of Old-Age and Survivors Insurance Wage Records and the Post-Enumeration Survey

B. J. MANDEL, IRWIN WOLKSTEIN, AND MARIE M. DELANEY

BUREAU OF OLD-AGE AND SURVIVORS INSURANCE,  
DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

## *Introduction*

The coordination of the Bureau of the Census's Post-Enumeration Survey (PES) responses with records of the Bureau of Old-Age and Survivors Insurance (OASI), like the other matching studies discussed in this volume, was part of the 1950 census record check program which attempted to provide benchmarks for evaluation of verbal survey responses. The check discussed in this paper consisted of a comparison of data on the number of employers and amount of wages reported to the PES with those reported on 1949 OASI tax returns.

The hope was that recognized deficiencies in comparisons of data from different surveys (since response errors are a function of the different survey conditions, and also some biases are present in all verbal responses) could be evaluated concretely by an actual case-by-case comparison of the most accurate survey material (PES) with tax records required by law from employers. If these records are presumed nearly perfect, differences between the survey responses and the records would be a direct measure of error in the survey responses. Because of inherent differences in the underlying concepts and coverage of the data compared, the wage match was only for a limited group of workers (those with a single employer during the year and less than \$3,000 wages reported in both sources). The chief problems arising from the 1949 OASI data were incomplete coverage and limitation of the tax base to \$3,000.

These difficulties, making results of the check somewhat inconclusive, appeared in various combinations at each step in the matching process. Briefly, the steps were: search for an OASI record for each of the 12,000 individuals in the PES sample, and matching for identity; comparison for each matched individual of number of

Note: The opinions expressed in this article do not necessarily reflect the views of the Bureau of Old-Age and Survivors Insurance. The writers are indebted to Rena Levine Berman for her assistance in planning the tabulations.

## MATCHING AND QUALITY STUDIES

employers reported in each source, and matching single employer for identity; comparison for each single employer individual of annual wages reported in each source. Despite the various limitations of the findings, the procedures and results may be useful beyond the initial purpose of the study.

All users of income statistics would benefit from the improvement in quality of published data which can result from coordination of data from two types of source. When the 1960 Census of Population is taken, OASI coverage of annual taxable earnings will be nearly complete, and the taxable wage base will be higher than that of 1949. The value of future comparisons will then increase.

The statistical program and related research activities of the Bureau of OASI will benefit from use of census earnings records which are more complete, for many workers, and can also provide information not now available to it on characteristics of income recipients (such as relationship to head of family and occupation of earner). In studies of income maintenance, data are needed on the income of the aged; in coverage extension studies, data on earnings in noncovered employment are required to estimate the additional tax yield and ultimate benefit levels. Further, the Bureau of OASI has a general interest in family income levels and occupational earnings, and a specific interest in the income of self-employed persons.

Besides the obvious value of OASI earnings records to the Census Bureau (for studying problems of interview response, possibly improving the accuracy of its data), the benefit payment records should supply an additional means of improving the accuracy of individual income data. OASI benefits are becoming an increasingly important income component—nearly \$7½ billion in 1957.

### *Procedural and Technical Problems of Coordination*

Difficulties, already touched upon, of coordinating two such dissimilar sets of data are increased by errors arising in clerical, coding, and tabulating operations, which can usually be eliminated or controlled. But the Bureaus of the Census and of OASI previously had demonstrated the feasibility of coordinating their respective data on employment, industry, and geographic location reported by employers for identical establishments, in their joint publication of *County Business Patterns* statistics.

### MATCHING RECORDS FOR IDENTITY OF INDIVIDUALS

In this study, the Bureau of OASI received from the Census Bureau its PES records for a sample of 12,000 individuals aged fourteen and

## OASI AND PES

over. The only advance information was that 70 to 75 per cent of the population aged fourteen and over had social security account numbers as of January 1, 1950. Approximately 7,800 account numbers were located by name in the OASI records, about two-thirds of the sample. (Because this ratio was calculated from the unweighted total, it cannot be automatically applied to the total population aged fourteen and over.) Possibly an additional 5 to 10 per cent had social security numbers but were not matched. In our judgment, the incompleteness of the match does not seriously detract from the validity of the findings.

Each case located in OASI records was classified as either a positive or probable match depending on the amount of identical information (date of birth, race, sex) in the two records.

### MATCHING EMPLOYERS IN INDIVIDUALLY MATCHED CASES

For each matched case, OASI figures on 1949 annual wages reported by each employer were compared with PES figures to secure matches in the number of employers and annual wages for that year. In addition to a comparison of number of employers reported, an attempt was made to match individual employers. (The employer-matching operation was confined to the major employer on each of the records, that is, the employer that paid the worker the largest amount of wages or salary in the year.)

The case records with matches for major employers were analyzed further to pinpoint discrepancies between PES and OASI records. The resulting classification consisted of five groups with variations in reporting major employer, and twelve subgroups containing variants in wages reported in the two sources (less than \$100 difference; \$100 or more difference).<sup>1</sup> Presumably, a difference in coverage

- <sup>1</sup> 1. Same major employer in PES and OASI records:
  - a. Less than \$100 difference in wage
  - b. \$100 or more or n.a. (not ascertainable) difference in wage
2. Employers reported as major to one agency and as nonmajor to the other:
  - a. Less than \$100 difference in wage between PES and OASI for each given employer
  - b. \$100 or more or n.a. difference in wage between PES and OASI for each given employer
  - c. Less than \$100 difference in wage between PES and OASI for PES major employer, and \$100 or more n.a. difference for OASI major employer
  - d. \$100 or more or n.a. difference in wage between PES and OASI for PES major employer, and less than \$100 difference for OASI major employer
3. OASI major employer not reported in PES, and PES major employer reported as nonmajor OASI employer:
  - a. Less than \$100 difference in wage of PES major employer
  - b. \$100 or more or n.a. difference in wage of PES major employer
4. PES major employer not reported in OASI, and OASI major employer reported as nonmajor in PES:
  - a. Less than \$100 difference in wage of OASI major employer

## MATCHING AND QUALITY STUDIES

accounts for a failure to match a major employer in the census record to the OASI record, while a failure to match a major employer in the OASI record can probably be traced to incomplete response to the PES. But other possibilities must be considered, for example, errors caused by overstatement and nonmatches due to incomplete identifying information, compensated—to an extent difficult to determine—by understatements and false matches.

### WAGE MATCHING

Comparison of wages reported in the two sources was necessarily preceded by the employer-matching process just described, because coordination of earnings data obviously could be carried out only when the same job was reported to PES and OASI. Since only about 2,000 individuals, or 30 per cent of the positive matches showed one and the same employer in the two records, the comparative study of total annual earnings had to be confined to this limited group.

#### *Results of Coordination Procedure*

The nature of the basic data did not permit a refined quantitative analysis of the accuracy of earnings reports in either source, but a number of qualitative facts became apparent from the comparison of number of employers reported in each source.

### EMPLOYER MATCHING

The following table shows how many employers were reported for the approximately 7,800 positively matched workers in the PES and OASI records for 1949:

NUMBER OF EMPLOYERS	PES <i>Records</i>	OASI <i>Records</i>	<i>Same Number in Both</i>
	<i>(per cent of positively matched workers)</i>		
Total	100.0	100.0	68.4
None	32.8	48.8	29.7
One	58.9	40.7	35.7
Two	7.1	7.8	2.7
Three	1.1	1.6	0.2
Four	0.1	0.7	<sup>a</sup>
Five and over	<sup>a</sup>	0.5	0

<sup>a</sup> Less than 0.05 per cent.

In about 68 per cent of the cases the same number of employers was reported in each source. Allowing for compensating errors, em-

- b. \$100 or more or n.a. difference in wage of OASI major employer
5. Other:  
 OASI and PES major employers not reported in PES and OASI respectively  
 No employer in either PES or OASI or both

## OASI AND PES

ployer matching even for the number of employers was probably imperfect in more than a third of the cases. Somewhat surprisingly, the PES records show fewer workers having two or more employers than the OASI records—8.3 per cent compared with 10.6 per cent. Noncovered employment undoubtedly accounts for most of the cases where PES records show more employers than OASI records; presumably response errors account for the fewer cases where the reverse is true. But even when there were more employers in the PES records, the major OASI employer was frequently not among them: <sup>2</sup>

EMPLOYER AGREEMENT	Per Cent of Pos- itively Matched Workers	Number of Employers Reported:		
		PES < OASI	PES = OASI	PES > OASI
Total	100	10	68	22
Same major employer	38	5	85	11
Major and minor reversed	1	39	47	14
OASI major not in PES, PES major a nonmajor in OASI	1	75	25	0
PES major not in OASI, OASI major a nonmajor in PES	5	4	10	86
OASI or PES major employer not in other record	8	22	70	8
None in either or both	53	6	58	36

Nearly 80 per cent of the positively matched records for women agreed on the number of employers, only about 60 per cent of the records for men:

SEX	Number of Employers Reported:		
	PES < OASI	PES = OASI	PES > OASI
Both sexes	9.5	68.4	22.1
Male	11.6	60.7	27.7
Female	6.8	78.6	14.6

This difference may be explained, in part, by the fact that women change jobs less often than men do, and therefore have fewer jobs to remember.

### WAGE MATCHING

In addition to the limitation of data for employer matching to individuals in the PES who were also (in 1949) covered by OASI, further limitation of this data for use in wage matching was required by the need to prove that the same wages from the same employer were to be compared in the two sources. This reduced our sample for wage matching to 30 per cent of the 7,800 positive matches, or

<sup>2</sup> The complexity of the code may have produced an excessive number of clerical errors which may account for some of the inconsistencies shown.

## MATCHING AND QUALITY STUDIES

about 2,240. While this group is obviously not a random sample of all 1949 wage earners, it is, in our opinion, not biased in any obvious manner for the purpose of measuring differences between wages reported in the two records for a given job. Errors due to failure to report an employer cannot, of course, be determined from an analysis of the data for this group.

The following table shows the percentage distribution of the records for positive matches with a single employer, by the amount of 1949 wages recorded in each of the two sources:

WAGES	PES Records	OASI Records	<i>Same in Both</i>
<i>(per cent of all workers)</i>			
Total	100.0	100.0	82.3
\$ 1-\$ 499	9.0	10.2	8.1
500- 999	7.4	7.0	5.1
1,000- 1,499	8.1	9.0	6.0
1,500- 1,999	11.1	11.0	7.9
2,000- 2,999	27.8	27.6	22.5
3,000 and over	36.6	35.2	32.7

Differences in distribution by earnings shown in the two sources are not significant in spite of different wage classes recorded in the two sources for nearly 18 per cent of the group. Obviously, compensating errors in reporting were responsible for the consistency in distributions. If real differences existed, they were in the direction of higher wages reported for the same employer in PES than in OASI. But, since the employer comparison showed a tendency on the part of PES respondents to omit second or third employers, underreporting in the survey of number of employers is partly compensated for by overreporting of annual earnings from a single job. This analysis of "bias" in survey response was made on the assumption that the OASI record could be used as a standard for comparison.

The analysis of wage differences was subject to the unavoidable restriction of 1949 OASI coverage—and therefore reported earnings—to \$3,000 per individual. In the following table, showing the percentage distribution of single employer matched records (about 2,240) by broad wages class and sex, the first line gives the relative size of the class reporting less than \$3,000 in each source, 61 per cent of the employer matched records.

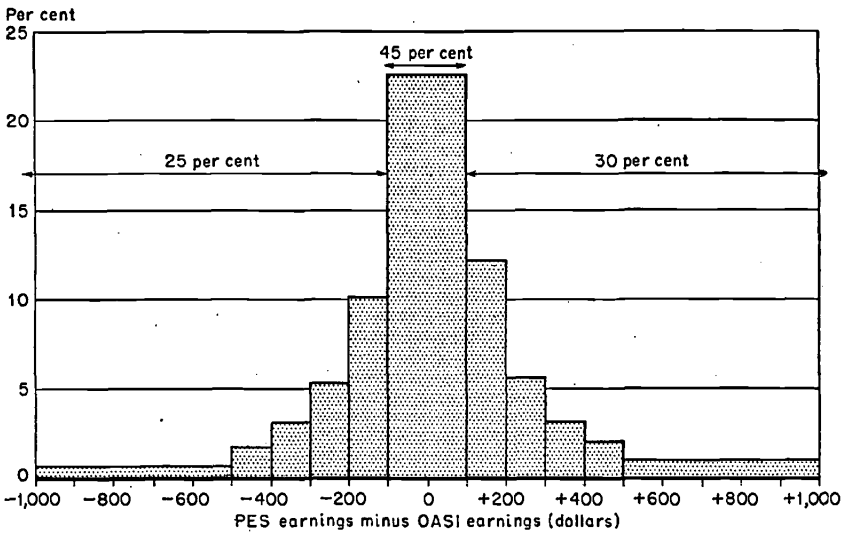
WAGES REPORTED		<i>Percentage Distribution</i>		
OASI	PES	Total	Male	Female
Less than \$3,000	Less than \$3,000	61	45	92
Less than \$3,000	\$3,000 or more	4	5	2
\$3,000 or more	Less than \$3,000	3	3	1
\$3,000 or more	\$3,000 or more	33	46	5

## OASI AND PES

If the two classes where less than \$3,000 was reported in one source and more than \$3,000 in the other source were considered in calculating numerical differences between wages reported in each source, without also including the class reporting more than \$3,000 in both sources, a biased presentation would result. Therefore, the analysis of wage differences was restricted to the class reporting less than \$3,000 to both PES and OASI, consisting of 61 per cent of the workers in the original wage matching group or less than 1,400. Chart 1 shows the percentage distribution of these workers by the amount of difference in reported wages.

### Chart 1

**Percentage Distribution of All Persons Having One Employer in PES and OASI, with Agreement on That Employer and on Wage Income under \$3,000, by Wage and Salary Difference, 1949**



Note: 1.1 per cent were in the minus \$1,000 or less class, 2.9 in the plus \$1,000 or more class.

This very symmetrical chart supports our previous findings, that is, the existence of a tendency for wage reporting errors to be reciprocally compensating, and our observation that, if any bias does exist for these matched single employer reports, it is in the direction of a slightly higher wage report in the PES. In the analysis of the basic data, tests for differences between reports of men and women and in reports of workers over and under age sixty-five showed that differences associated with sex or age were insufficient to significantly influence the distribution.



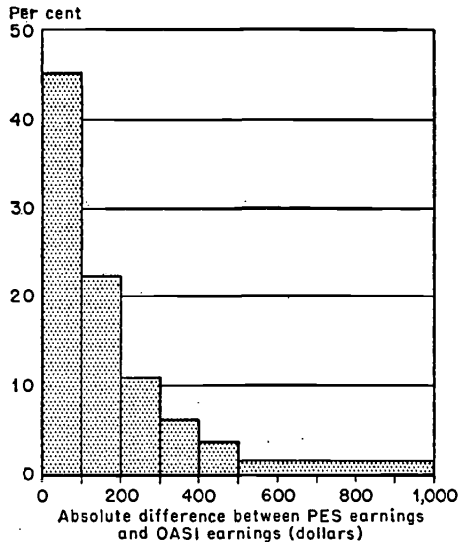
## MATCHING AND QUALITY STUDIES

Chart 2 shows the percentage distribution of workers by amount of absolute difference in wages regardless of the direction of the difference.

About 45 per cent of the records for workers with earnings of less than \$3,000 in both sources differed by less than \$100, and two-thirds by less than \$200. For less than 5 per cent of the workers

Chart 2

Percentage Distribution of All Persons Having One Employer in PES and OASI, with Agreement on That Employer and on Wage Income under \$3,000, by Absolute Wage and Salary Difference, 1949



Note: 4.0 per cent were in the \$1,000 or more class.

in the group was the difference as much as \$1,000. A rough calculation was made of the mean squared difference between wages reported by the two sources, and it was compared with the variance of the wage distribution. For the limited area of comparison, the mean square response error appears to be much less than the variance of the wage distribution.

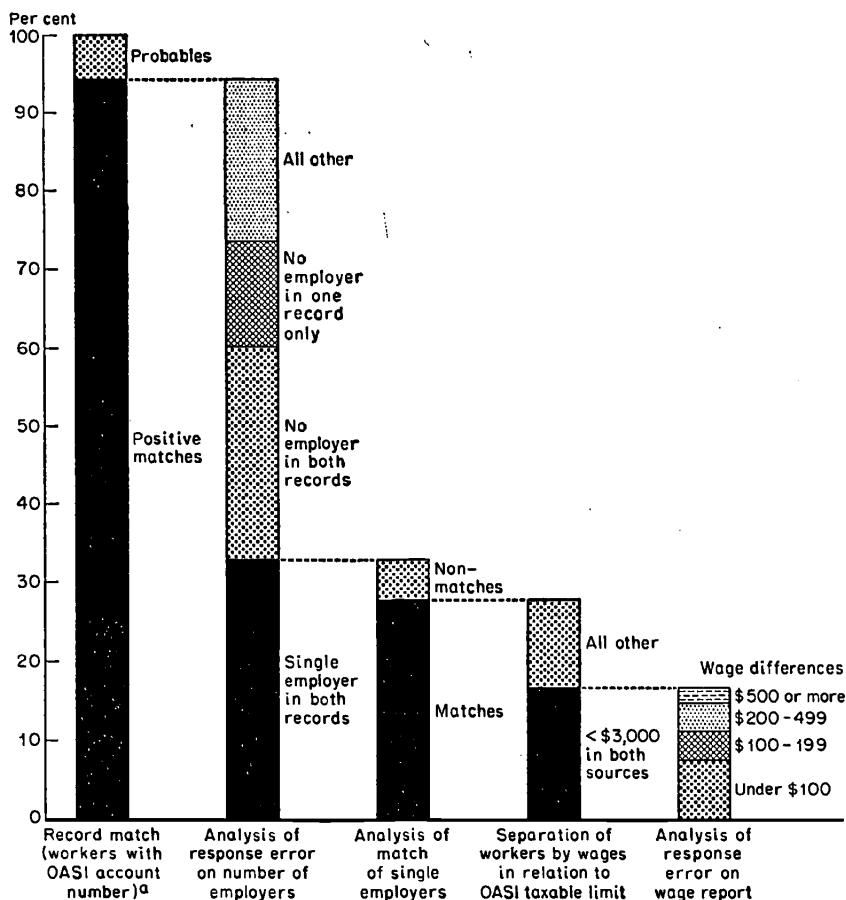
### *Summary and Conclusions*

Chart 3 traces in terms of percentages the number of persons originally identified in the OASI-PES sample through the various stages of analysis.

It is apparent that with successive steps in the analysis, the orig-

Chart 3

Percentage of Workers with OASI Account Numbers Used in Successive Stages of Analysis, 1949



<sup>a</sup> Represents 65 per cent of all workers in PES sample.

inally identified group of 7,800 cases was gradually reduced. These reductions were made because of limitations in scope of operations and the limitations in 1949 of old-age and survivors insurance coverage to only about 70 to 75 per cent of all jobs and the first \$3,000 of earnings. Despite the limited nature of the comparisons possible in this study, we believe that the following generalizations can be made:

1. The belief of the Census Bureau that incidental employment is frequently forgotten is confirmed by the finding that some re-

## MATCHING AND QUALITY STUDIES

spondents fail to report additional employers if they have more than one job in a year.

2. Amounts reported do not appear to be significantly biased, at least not for workers with less than \$3,000 a year.

3. The errors found in PES reported wages which could be compared to those in OASI appear to be small relative to the variance of the wage distribution. For samples small in comparison with the population, the contribution of reporting error to total error of estimate of mean wages for the groups studied is probably negligible, compared with sampling error.

4. For most individuals with records in both census and OASI sources, matches can be achieved with a reasonable expenditure of effort, provided proper identifying information is in the census record. The limited effort devoted to matching in the present study resulted in matching the records for about 90 per cent of the covered persons.

It is appropriate to add some recommendations about the future of coordination of household interview responses with OASI records. We have no doubt that such coordination has value in filling gaps in knowledge about workers and beneficiaries under the old-age and survivors insurance program. As far as the interests of Bureau of the Census and users of census statistics are concerned, accuracy checks and response research are unquestionably necessary and will, therefore, continue. The various checks already made, however, should be reviewed critically, and the combination of sources providing the "best" results with the least expenditure should be chosen. In making this selection the scope of the OASI program by the time of the proposed check will be an added favorable factor, for the four increases in coverage since 1949 will probably be expanded, making results of future wage coordination studies more significant than those obtained from the 1949 study. Furthermore, difficulties in matching procedures may be lessened; with nearly universal coverage of workers under OASI, the necessity for matching employers may be eliminated. One factor to be stressed in the effectiveness of a matching study is the availability of sufficient identifying information, which depends largely on efforts expended to collect it in the enumeration process.