Chapter 6.-Panel Studies of Earnings Histories

1. Purposes of These Studies

The major data analysis project undertaken by the Panel was classification and modeling of earnings histories of workers. In this project,the Panel benefited from many useful ideas given us by members of the technical staff of the Social Security Administration, who also made most of the statistical tabulations for us.

This project had three basic goals. The first was to test the validity of the assumption of constant exponential growth in earnings that has strongly influenced benefit design up to the present. The Panel undertook to examine, by statistical analysis of workers' lifetime wage patterns, how the facts conform to the constant exponential growth rate pattern. Findings from such an analysis can significantly influence benefit design.

The second goal was to develop a pool of statistics and a simulation model that could be used to test alternative benefit formulas.

The benefit formulas examined by the Panel were of two broad types. there were formulas based on short averaging periods, e.g., those based on a worker's highest five or ten years of earnings. Second, there were formulas based on longer averaging periods but with past earnings equalized through an indexing process. To test these formulas, extensive earnings data are needed.

A question about formulas of the first kind is whether the years used for benefit determination are close enough to retirement to reflect consistently the income to be replaced. Distortions arise when periods of highest earnings are close to retirement for some people but distant from retirement for others; workers with identical *real* earnings are likely to get considerably different benefits, those whose highest earnings are closest to retirement being the most generously treated. Consequently it is important to be informed of the distributions of these high earnings years.

In comparing methods for indexing money earnings, the correlation between the average indexed earnings and the average (unindexed) earnings used in present law is of significance. If it is positive and high, it is relatively easy to construct a benefit formula that will produce benefit amounts consistently close to those of present law.

A more important consideration in appropriately adjusting earnings histories by an indexing process is the applicability of the index to various groups of workers. We need to know to what extent using an average rate of wage increase overstates the wage increases for some groups and understates them for others. Hence, the variability of earnings increases needs investigating.

The third goal of the project was to develop methods by which more realistic statistically based earnings histories may be suggested for use in official cost estimation procedures in place of the present official simulation technique.

2. Sources of Data

Much is already known about the distribution of earnings in our national work force. However, most of these figures are cross-sectional, i.e., they tell us about earnings distributions at one particular time. For social security work earnings patterns of individuals over long periods are needed.[1]

Two compilations of data from the ongoing research activity of the Social Security Administration were made available to us. The first was the *0.1 percent Continuous Work History Sample (CWHS).* This sample of random social security numbers provides chronological information on one out of each thousand

[1] Because of absence of **sufficient such information, the benefit illustrations in Chapter 3 were constructed principally from census data.**

workers with wage histories. The information includes annual taxable earnings and estimated annual total earnings in covered employment for the period 1956-1972. The Panel's work would have been facilitated if a longer continuous record had been available. Information about these seventeen years, years of relative economic stability but with some cycles although no deep economic dislocations, was indeed useful. However, the earnings fluctuations revealed by this sample undoubtedly understate the fluctuations that would mark a major recession.

The second source of data consisted of two random samples of benefit awards made in 1974: the first, 8,399 awards spread over July-December; the second, 3,501 awards in December. These samples were helpful in confirming results derived from the CWHS. The tables and charts in Appendix A all come from CWHS figures.

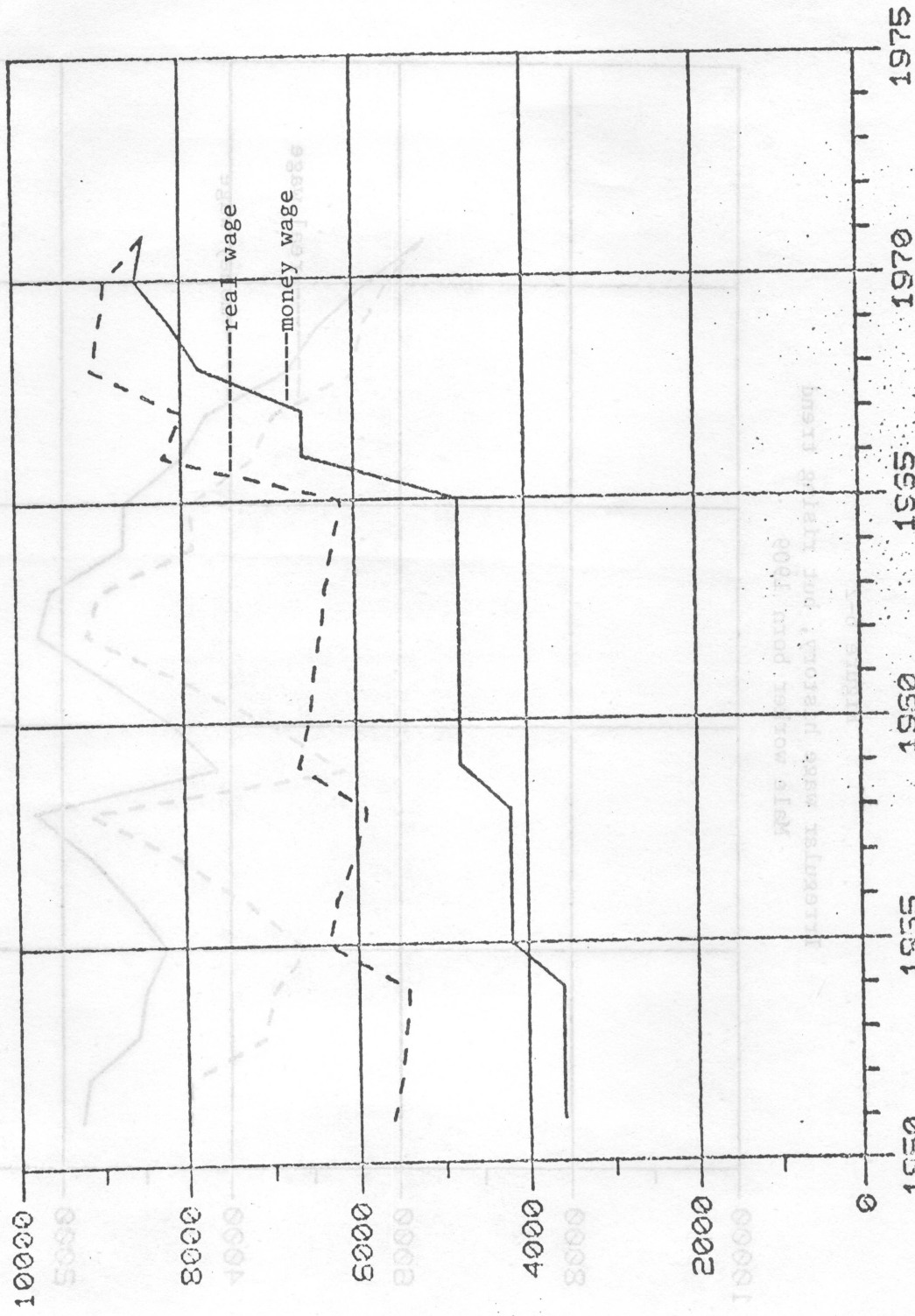
The CWHS figures and the samples unfortunately give limited information about the people in the sample. Our classifications and models were, on this account, necessarily based on age, sex, and wage history only. Additional particulars would have been analytically useful.

3. Derived Statistics

The data were used to develop graphical displays and tabulations that led to further classifications and modeling. Computer-produced graphs of wages subject to payroll tax and of real wages (adjusted by CPI) for workers in the CWHS file born in 1909, 1919, and 1924 were studied. A crude review of these graphs (see samples in Figures 6-1, 6-2, 6-3, and 6-4) revealed that about half of the male workers enjoy steadily increasing wage trends. A more important conclusion was that wage histories suffer from great variability.

Figure 6-1

Steady Rising Pattern Male Worker Wage History:
Born 1909, Consistently at the Taxable Maximum



70-577 O - 76 - 4

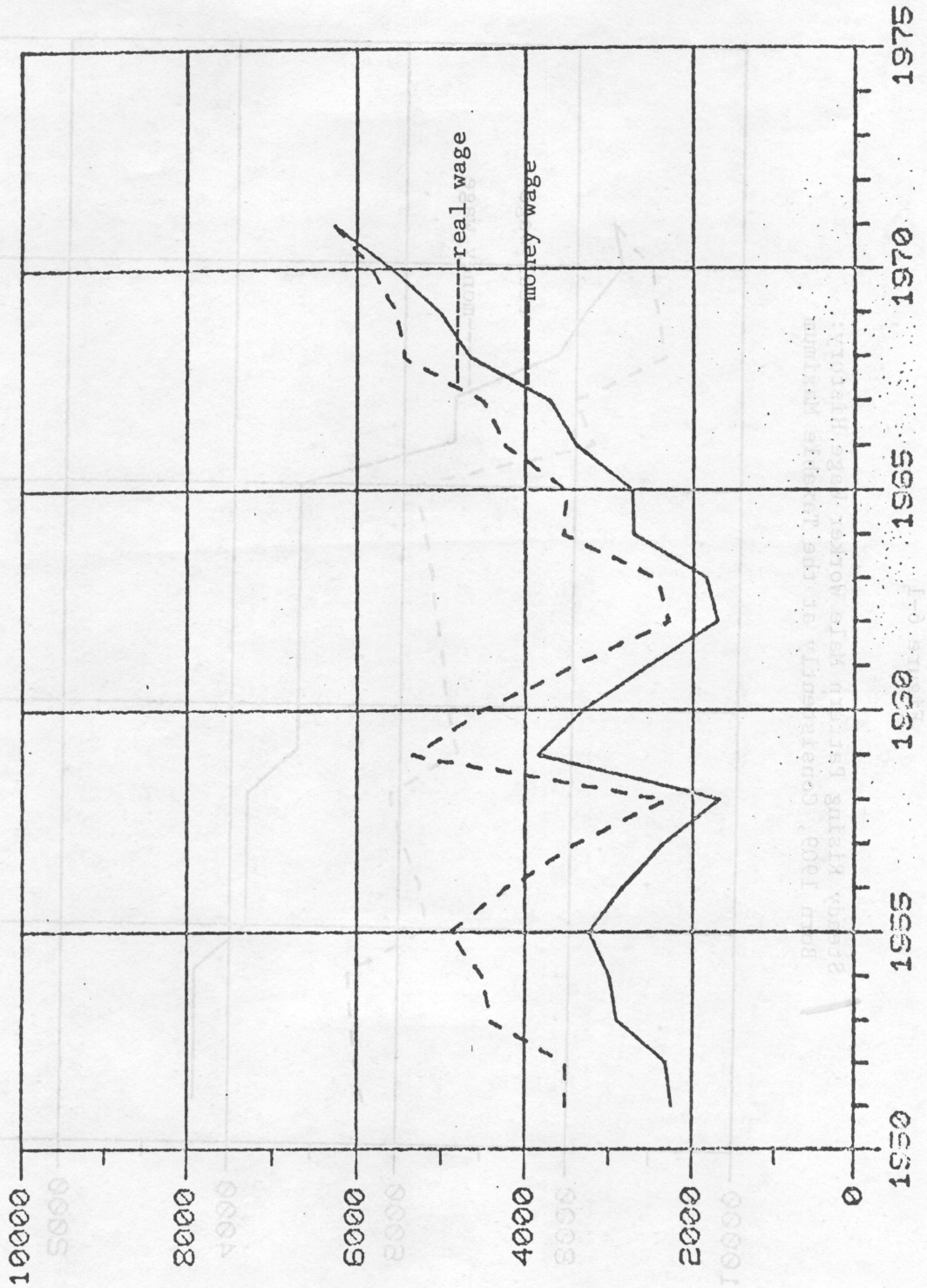


Figure 6-2

Irregular wage history, but rising trend
Male worker born 1909

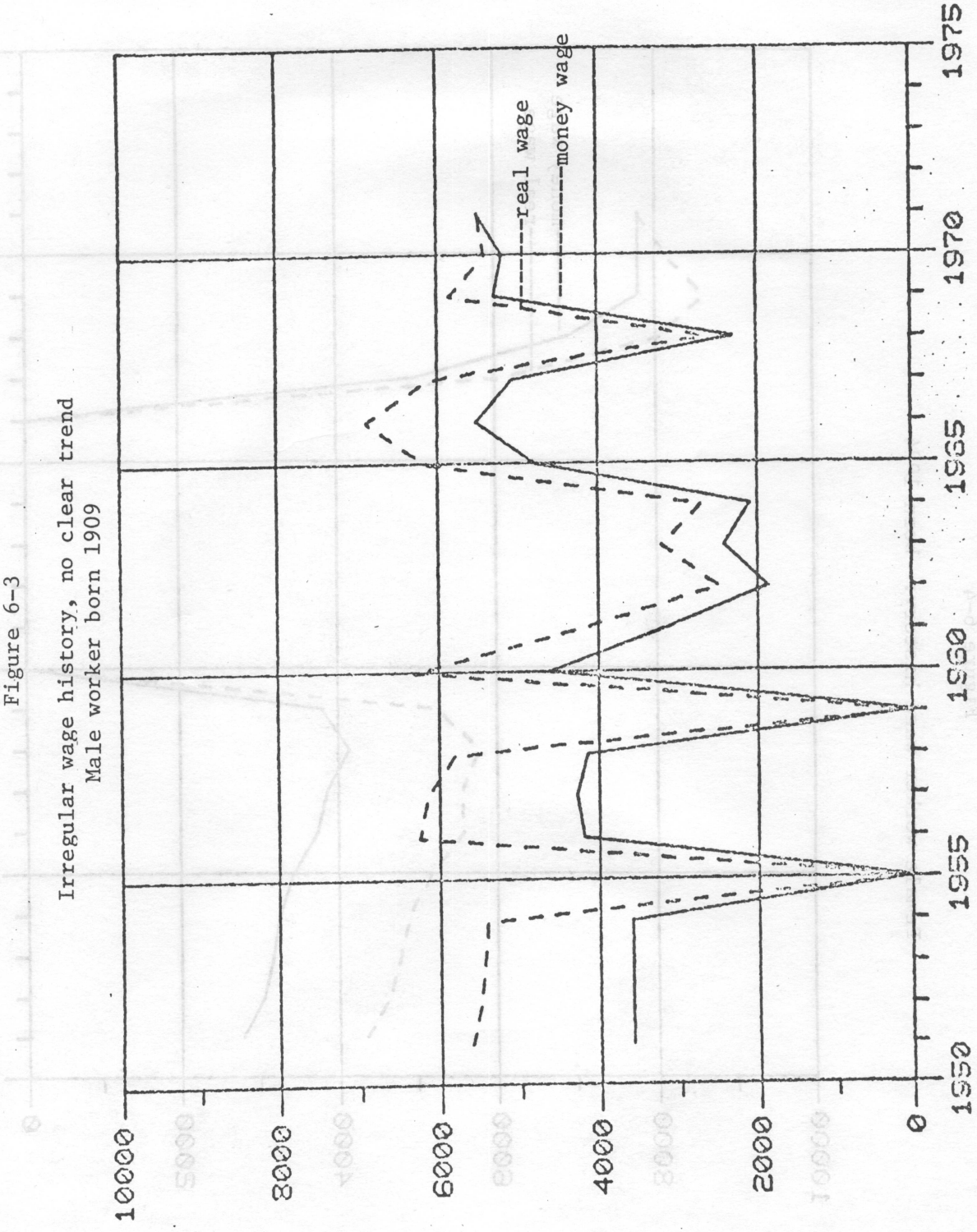


Figure 6-3
Irregular wage history, no clear trend
Male worker born 1909

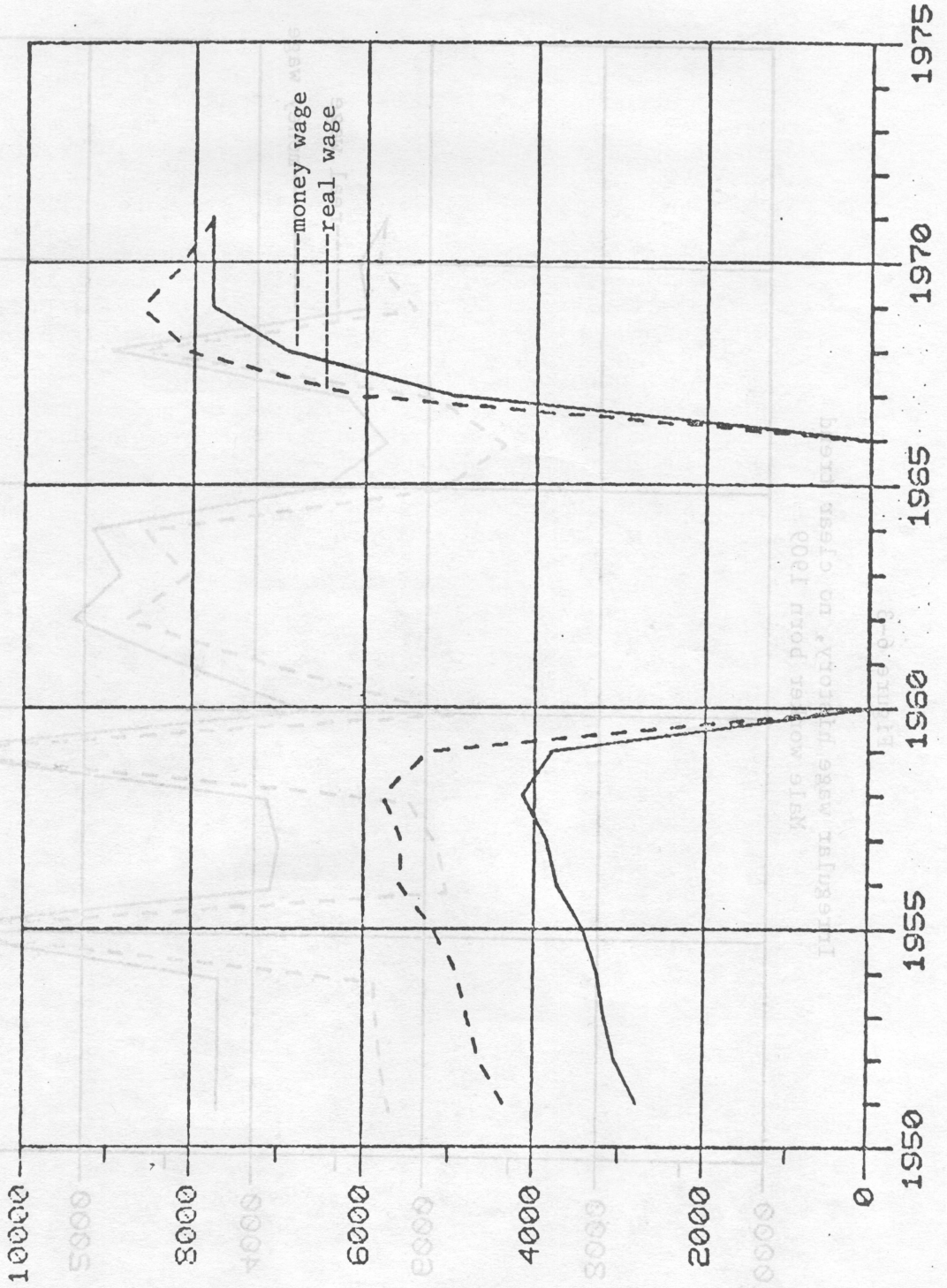


Figure 6-4

Female worker wage history: born 1924

Next, a series, as follows, of six basic types of analysis were made:

a. Annual rates of change in total estimated earnings (Appendix A, Section 1) for a sample from birth cohorts 1907, 1917, 1927, 1937 were tabulated. The average annual rates of change through 1957-1972 within these cohorts (excluding workers with zero earnings and extraordinary earnings changes) and their standard deviations were calculated. The principal conclusions are that rates of changes are higher for younger workers than for older workers and that there is much variation in wage change rates, both for specific people from year to year and among different people. See Figure 6-5 for an example of this analysis.

b. Analysis, with similar exclusions, was made of rates of change in estimated total earnings of workers who were persistently within low, middle, and high earnings groups. The principal conclusions from the prior analysis were reinforced. It appears also that workers with higher earnings have larger average rates of increase. Figure 6-6 summarizes the results for average rates of earnings change for workers persistently in the lowest one-third, the middle, and the highest one-third of earnings within year of birth and sex groups. This complete analysis appears in Appendix A, Section 1.

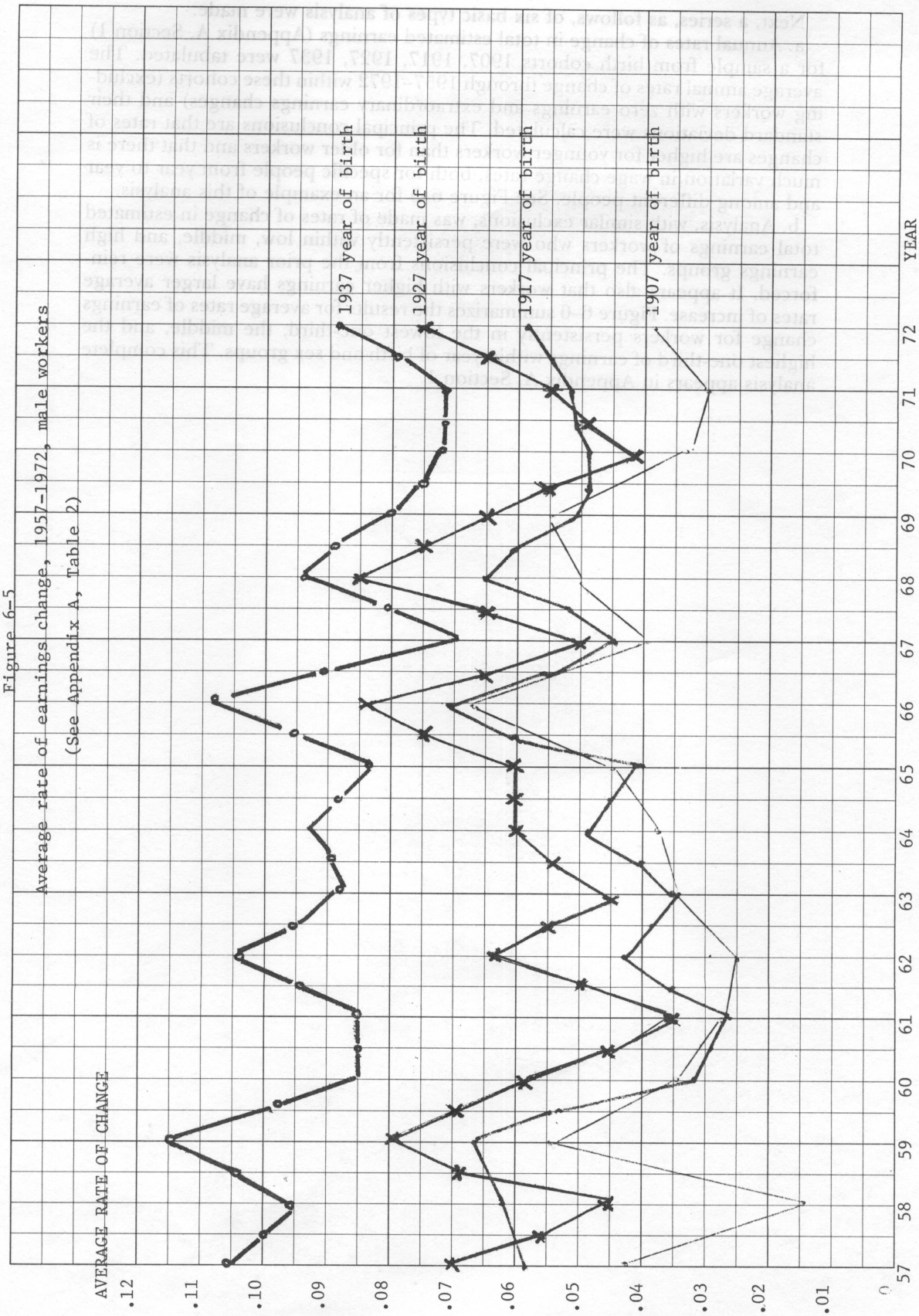


Figure 6-5
Average rate of earnings change, 1957-1972, male workers
(See Appendix A, Table 2)
AVERAGE RATE OF CHANGE
.12
.11
.10
.09
.08
.07
.06
.05
.04
.03
.02
.01
0
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
YEAR
1937 year of birth
1927 year of birth
1917 year of birth
1907 year of birth

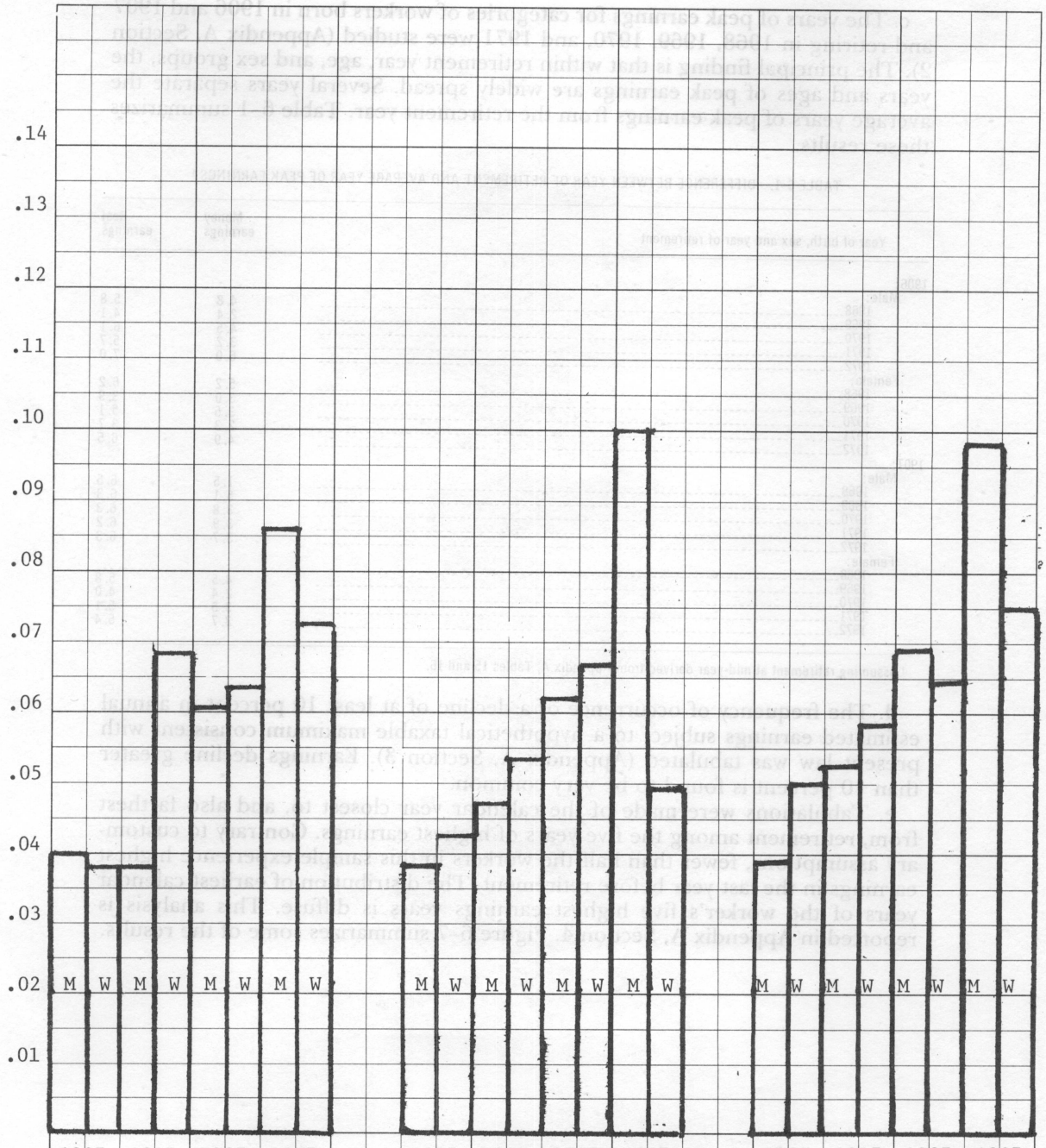


Figure 6-6

Average rate of earnings change 1956-1972,
for workers persistently in the same earnings level class
(See Appendix A, Table 4-14)

c. The years of peak earnings for categories of workers born in 1906 and 1907 and retiring in 1968, 1969, 1970, and 1971 were studied (Appendix A, Section 2). The principal finding is that within retirement year, age, and sex groups, the years and ages of peak earnings are widely spread. Several years separate the average years of peak earnings from the retirement year. Table 6-1 summarizes these results.

TABLE 6-1.—DIFFERENCE BETWEEN YEAR OF RETIREMENT AND AVERAGE YEAR OF PEAK EARNINGS [1]

Year of birth, sex and year of retirement	Money earnings	Real earnings
1906:		
Male:		
1968	4.8	5.8
1969	2.4	4.1
1970	4.5	6.1
1971	3.2	5.7
1972	6.0	7.0
Female:		
1968	5.2	6.2
1969	3.0	3.9
1970	3.5	5.1
1971	2.3	3.7
1972	4.9	6.5
1907:		
Male:		
1968	1.5	6.5
1969	5.1	6.3
1970	3.8	6.3
1971	3.9	6.2
1972	3.7	6.3
Female:		
1968		
1969	4.5	5.8
1970	2.4	4.0
1971	3.5	5.1
1972	3.7	5.4

[1] Assuming retirement at mid-year derived from Appendix A, Tables 15 and 16.

d. The frequency of occurrence of a decline of at least 10 percent in annual estimated earnings subject to a hypothetical taxable maximum consistent with present law was tabulated (Appendix A, Section 3). Earnings decline greater than 10 percent is found to be very common.

e. Tabulations were made of the calendar year closest to, and also farthest from, retirement among the five years of highest earnings. Contrary to customary assumptions, fewer than half the workers in this sample experience highest earnings in the last year before retirement. The distribution of earliest calendar years of the worker's five highest earnings years is diffuse. This analysis is reported in Appendix A, Section 4. Figure 6-7 summarizes some of the results.

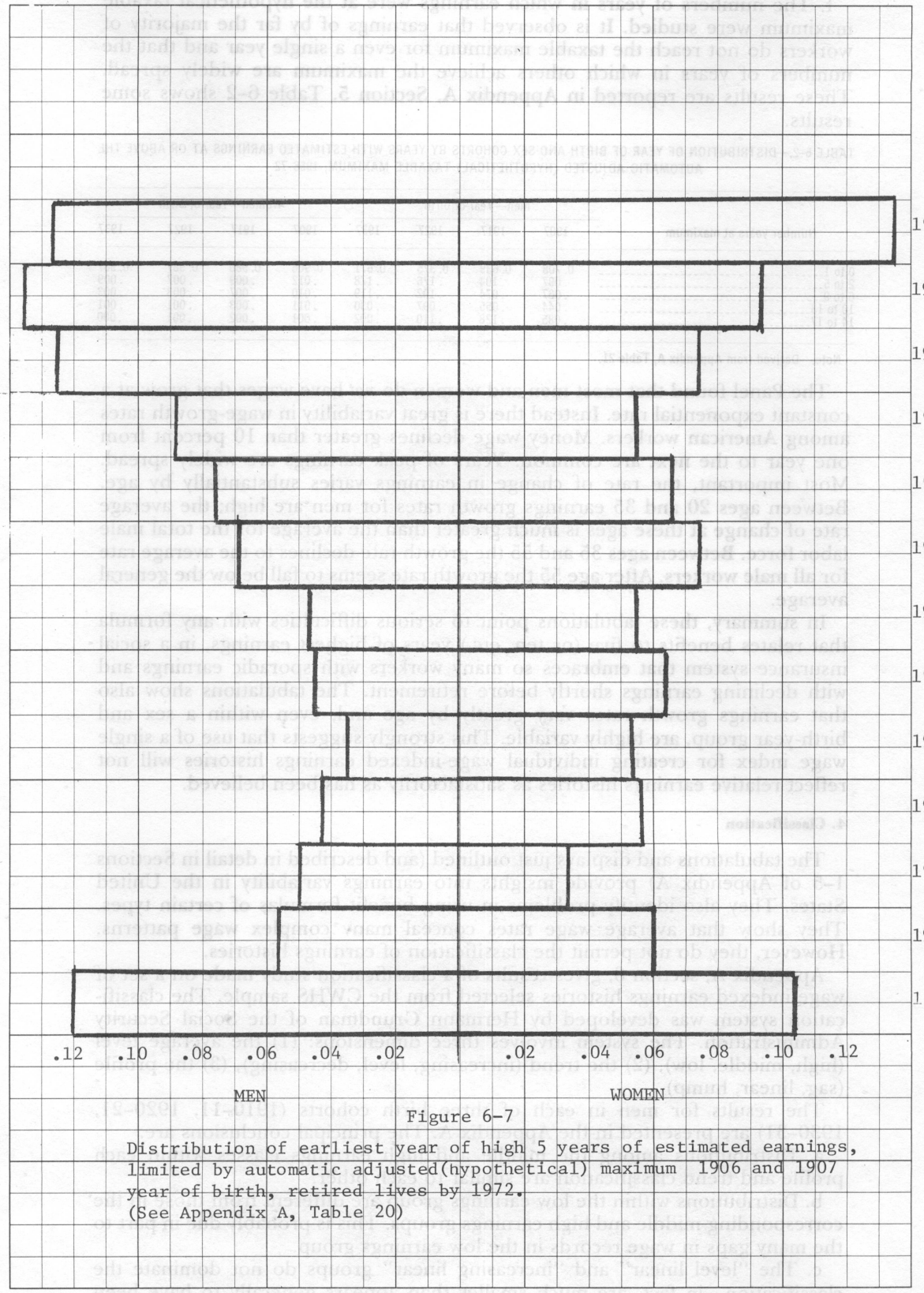


Figure 6-7

Distribution of earliest year of high 5 years of estimated earnings, limited by automatic adjusted(hypothetical) maximum 1906 and 1907 year of birth, retired lives by 1972.
(See Appendix A, Table 20)

f. The numbers of years in which earnings were at the hypothetical taxable maximum were studied. It is observed that earnings of by far the majority of workers do not reach the taxable maximum for even a single year and that the numbers of years in which others achieve the maximum are widely spread. These results are reported in Appendix A, Section 5. Table 6-2 shows some results.

TABLE 6-2.-DISTRIBUTION OF YEAR OF BIRTH AND SEX COHORTS BY YEARS WITH ESTIMATED EARNINGS AT OR ABOVE THE AUTOMATIC ADJUSTED (HYPOTHETICAL) TAXABLE MAXIMUM, 1956-72

Number years at maximum	Men-Year of birth				Women-Year of birth			
	1907	1917	1927	1937	1907	1917	1927	1937
Otol	0.708	0.639	0.375	0.671	0.966	0.980	0.987	0.987
2 to 5	. 097	. 104	. 176	. 178	. 012	. 009	. 007	. 009
6 to 9	. 067	. 074	. 102	. 119	. 010	. 007	. 003	. 003
10 to 13	. 044	. 055	. 097	. 030	. 011	. 003	. 001	. 001
14 to 17	. 085	. 128	. 110	. 002	. 001	. 002	. 001	.000

Note.-Derived from Appendix A, Table 21.

The Panel found that most men and women do *not* have wages that grow at a constant exponential rate. Instead there is great variability in wage-growth rates among American workers. Money wage declines greater than 10 percent from one year to the next are common. Years of peak earnings are widely spread. Most important, the rate of change in earnings varies substantially by age. Between ages 20 and 35 earnings growth rates for men are high; the average rate of change at these ages is much greater than the average for the total male labor force. Between ages 35 and 55 the growth rate declines to the average rate for all male workers. After age 55 the growth rate seems to fall below the general average.

In summary, these tabulations point to serious difficulties with any formula that relates benefits to five (or ten, etc.) years of highest earnings, in a social insurance system that embraces so many workers with sporadic earnings and with declining earnings shortly before retirement. The tabulations show also that earnings growth rates vary greatly by age and, even within a sex and birth-year group, are highly variable. This strongly suggests that use of a single wage index for creating individual wage-indexed earnings histories will not reflect relative earnings histories as satisfactorily as has been believed.

4. Classification

The tabulations and displays just outlined (and described in detail in Sections 1-5 of Appendix A) provide insights into earnings variability in the United States. They also identify problems in using benefit formulas of certain types. They show that average wage rates conceal many complex wage patterns. However, they do not permit the classification of earnings histories.

Appendix A, section 6, gives results of a classification study made on a set of wage-indexed earnings histories selected from the CWHS sample. The classification system was developed by Hermann Grundman of the Social Security Administration. The system involves three dimensions: (1) the average level (high, middle, low), (2) the trend (increasing, level, decreasing), (3) the profile (sag, linear, hump).

The results for men in each of three birth cohorts (1910-11, 1920-21, 1930-31) are presented in the Appendix A. The principal conclusions are:

a. Distributions among the middle and high earnings classes within each profile and trend classification are similar to each other.

b. Distributions within the low earnings group are different from those in the corresponding middle and high earnings groups. This is probably due in part to the many gaps in wage records in the low earnings group.

c. The "level linear" and "increasing linear" groups do not dominate the classification-in fact, are much smaller than appears generally to have been believed.

d. The youthful cohort shows the highest proportion of people whose wage-indexed earnings trend upward.

5. Correlations Among Wage Histories

Each of the benefit formulas considered by the Panel involves some form of earnings averaging. Differences among them are in the lengths of the averaging period and the weights employed. To gauge the problems of transition from one benefit formula to another, the equivalence between average earnings computed under present law, under a price-indexed formula, under a wage-indexed formula and under High-10 and High-5 formulas, were studied. For this the earnings histories in a random CWHS sample of 4,320 workers were used.

The coefficient of correlation used in table 6-3 that follows is the normal statistical measure of relationship such that a value of 1.0000 marks perfect linear relationship and a value of 0.0000 shows that the items compared are varying completely independently. Increasingly negative values portray inverse relationships.

TABLE 6-3.-COEFFICIENTS OF CORRELATION BETWEEN THE AVERAGE MONTHLY WAGE CALCULATED UNDER VARIOUS ALTERNATIVE BENEFIT FORMULA-TOTAL (4,320 LIVES)

	Price indexed	Wage indexed	High 10	High 5
Current law	0.9947	0.9794	0.9739	0.9083
Price indexed		. 9943	. 9605	. 8819
Wage indexed			. 9363	. 8476
High 10		-----		. 9672

These results lead to the conclusion that a price-indexed formula produces results more closely correlated with the present law than the other systems do.

6. Regression Model

To fill the need that exists for simulating earnings histories, experiments were conducted with several statistical models developed from CWHS data. Although the CWHS lacks information on many useful variables, such as occupation and education, which others have found helpful, the earnings, age and sex data were available to develop parameters. Regression models were developed for male lives using the logarithm of wage-indexed earnings as the response variable and age and individual dummies as independent variables.

Details of this modeling effort appear in Appendix B. The typical lifetime profile of wage-indexed earnings that emerged is shown in Figure 6-8.

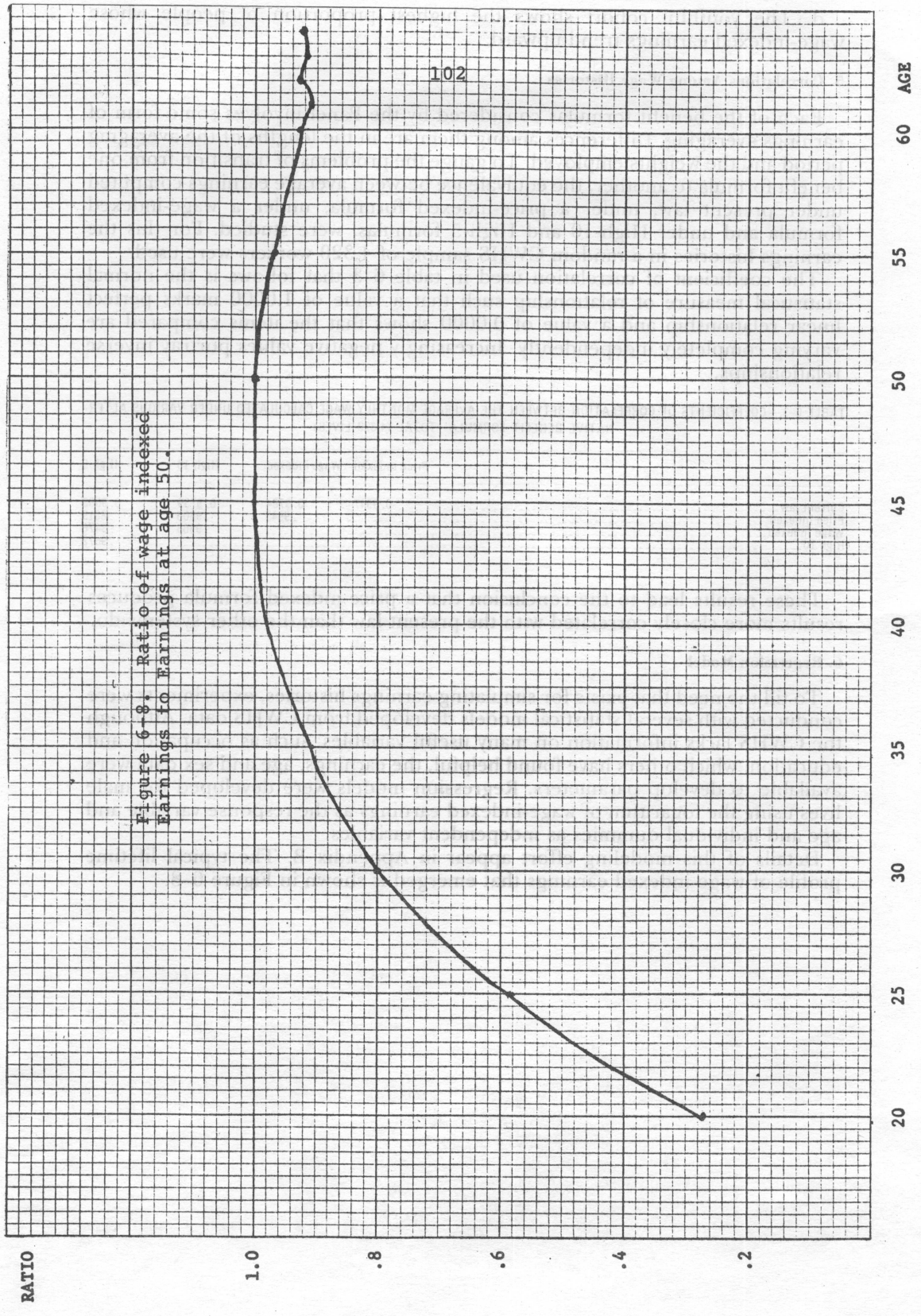


Figure 6-8. Ratio of wage indexed Earnings to Earnings at age 50.

In addition, a simple probability model for simulating the occurrence of zero earnings and return from the zero earnings state was constructed.

The principal conclusions drawn from this whole project are:

1. Rates of change in earnings are high at ages 20 to 35.
2. Rates of change in wage-indexed earnings are not large beyond age 40.
3. There is much variability, particularly for low earners, around the response variable (earnings).
4. The probability of zero covered earnings in the year that follows any year of positive earnings proves to be 3 percent. This is consistent with the finding that approximately 60 percent of men in the sample had a full 16 years of positive earning in the 16 years examined.
5. The random influences upon wage growth have important bearing upon the effect of lengthening the average period in benefit computation. (This lengthening will occur under present law or under either of the indexing proposals under consideration.)