

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

Volume Title: The Effect of Pension Plans on Aggregate Saving: Evidence from a Sample Survey

Volume Author/Editor: Philip Cagan

Volume Publisher: NBER

Volume ISBN: 0-87014-409-X

Volume URL: <http://www.nber.org/books/caga65-2>

Publication Date: 1965

Chapter Title: Description of the Data and Sources of Possible Error

Chapter Author: Philip Cagan

Chapter URL: <http://www.nber.org/chapters/c1671>

Chapter pages in book: (p. 8 - 19)

DESCRIPTION OF THE DATA AND SOURCES OF POSSIBLE ERROR

One way of determining the effect of pension plans on aggregate personal saving is to examine the trend of this kind of saving. Since the late 1940's, the percentage of personal income saved shows considerable variation from year to year, but has if anything declined slightly (even disregarding the early post-World War II years), while the percentage saved through group pension plans rose almost one point over the same period.¹ The trouble with these data, of course, is that numerous

¹ For all households, the percentage of personal income saved in recent years, through public and private pension plans and in all financial assets, is shown below. The denominator of these ratios is total personal income. If it were personal disposable income instead, the rate of saving would be about a percentage point higher. Numerators and denominators include employers' payments into private pension funds.

	Pension Saving Ratio	Total Personal Saving Ratio		Pension Saving Ratio	Total Personal Saving Ratio			
		SEC	Commerce		SEC	Commerce		
1945	1.3	8.0		1955	1.8	6.4	5.1	
1948			6.4	1956	1.9	7.3	6.2	
1949			4.5	1957	1.9	7.2	5.9	
1950			5.7	1958	2.1	7.5	6.2	
1951	1.6	7.1	6.8	1959	2.2	6.9	5.0	
1952			6.7	1960	2.1	5.2	4.3	
1953			6.4	1961	2.2	6.9	5.1	
1954			5.7	1962	2.1	6.8	4.9	
					1963	2.2	6.3	4.4
					1964			5.3

Source: Pension saving (change in total assets of all public and private pension funds): 1950-63, SEC, *Statistical Bulletin* (earlier unpublished data supplied by the Commission); 1945-49, private insured plans from Institute of Life Insurance, *Private and Public Pension Plans in the United States*, private noninsured and state and local government employee funds from Raymond W. Goldsmith, *A Study of Saving in the United States*, Vol. I, Table I-15, p. 468 and Table G-19, p. 1073, Railroad Retirement System and U.S. Civil Service Retirement System from annual reports of their respective boards adjusted to agree with 1945 and 1950 figures in *Private and Public Pension Plans in the United States*.

Personal saving: Securities and Exchange Commission data (with the increase in

factors besides pension growth affect aggregate saving. There is no way to adjust precisely (or even imprecisely) for the other effects, which conceal the relation between pension growth and other saving.

Another method is to compare, by means of surveys, the saving of households with and without pension plans, first adjusting for all other respects in which the households may differ. This is the approach followed here. To avoid the prohibitive cost of a survey large enough to be useful, the data were obtained through use of a survey already in operation studying consumer expectations and expenditures with the use of questionnaires sent periodically to the same list of members of Consumers Union of the United States (the national consumer product-testing organization). Questions were added on saving, pension coverage, and other related items to two of these questionnaires, one sent in the autumn of 1958 and another in the spring of 1959, each pertaining to the twelve months preceding the date of the survey. Their coverage therefore overlaps the six-month period preceding the autumn of 1958. Questions asked in the first survey provided preliminary information that suggested additions and improvements for the second. The second survey is technically superior in several respects, and it provides most of the data reported below, though some information from the first is also used. Also, two questions were taken from a questionnaire sent in the spring of 1958. The questions used from the three questionnaires will be found at the end of the Appendix.

1. Characteristics of the Survey Sample

Survey data are justifiably suspect for the accuracy of answers, the applicability of the sample to all households, and the period covered. Consumers Union members, as will be demonstrated, constitute a select group, and it is necessary to know how their saving differs from the national average.

First, a comment on the accuracy of answers. Written replies to questionnaires seldom lack inaccuracies, especially when, as in these surveys,

government pension funds excluding OASDI added): 1945-55, *Historical Statistics of the United States, Colonial Times to 1957*, Table F261, p. 153; 1956-63, SEC, *Statistical Bulletin*, July issues; Department of Commerce data, *Survey of Current Business*, August 1965, Table 5.

Personal income: *ibid.*

detailed questions deal with financial matters that require considerable sophistication and the consultation of records. Many respondents may have misinterpreted some questions, may not have consulted records (indeed, may not have had accurate records), and may have lost patience with the long questionnaire and made mistakes. The saving, wealth, and pension questions took time to answer, yet they were preceded on the questionnaires by others that took even longer. The data undoubtedly contain considerable reporting error (presumably mostly random but, as noted later, not always). As we shall see in the following section, the variance of the saving-income ratio among households is sizable.

Two factors ease the problem of reporting errors. One is the large sample, which numbered 15,873. Insofar as reporting errors are random, they tend to offset each other in a large sample, and the results may still be fairly accurate for the group as a whole. (Possible sources of bias from nonrandom errors will be noted later.) The second factor is the character of the households surveyed. They are above the national average in income and education, and therefore presumably in intelligence. They are also conscientious, as indicated by their membership in Consumers Union and by their willingness to fill out long and complicated questionnaires. The response rate was fairly high—nearly two-thirds—and the number of questions left blank surprisingly small. This degree of cooperation was not wholly accidental, however, since questionnaires were sent only to those among 100,000 Consumers Union members who on an earlier questionnaire had indicated their willingness to answer a special series of detailed questionnaires about their financial affairs and buying plans. What is gained in the greater accuracy of the responses, therefore, is partly lost in the select coverage of the sample. This was an advantageous trade, because the data, though subject to considerable error, are usable; if they had been elicited in the same way from a sample of all U.S. households, they would probably have been much less accurate and complete.

Is the sample nonetheless so select as to lack relevance to the economy at large? It is certainly unrepresentative of all households. Tables 1 and 2 compare the sample used in the subsequent analysis with all households in the United States for education and income. College graduates are predominant among the sample households and the

below-\$5,000 income classes are deficient, though the income distribution is more similar to that of total U.S. households if the very low class of less than \$3,000 is excluded. There are also a disproportionately large number of teachers and government workers and a small number of wage earners. Most of these distributional limitations may

TABLE 1

Comparison of Education: Consumers Union Sample and All U.S. Households, 1957-58

Education	Consumers Union Sample ^a	All U.S. Households ^b
High-school graduate or less	17.9	82.1
Some college	22.4	8.8
College graduate or more	59.7	9.2

^a Reduced sample used in the subsequent analysis (described later in this chapter) after excluding self- and not-employed, incomplete questionnaires, and households with unusual gains or losses over \$1,000 or with saving greater in absolute amount than 49 per cent of income.

^b Based on a sample survey of the labor force 18-64 years old in March 1957. *Statistical Abstract*, 1959, p. 109.

TABLE 2

Comparison of Income Levels: Consumers Union Sample and All U.S. Households, 1958-59

Income Level	Consumers Union Sample ^a	U.S. Households ^b	
		All	All Excluding Incomes Less Than \$3,000
Less than \$3,000	0.5	33.0	--
3,000 - 3,999	1.7	11.1	16.6
4,000 - 4,999	4.7	12.4	18.5
5,000 - 9,999	59.2	35.1	52.4
10,000 - 14,999	26.3	6.4	9.6
15,000 - 24,999	6.3	1.6	2.4
25,000 and over	1.2	0.4	0.6

^a The sample used in the subsequent analysis after exclusions (see note a, Table 1).

^b Based on a sample survey of families and unrelated individuals in 1959. Bureau of the Census, *Current Population Reports*, P-60, No. 33, January 15, 1960, Table 5.

be avoided by breaking down the sample by income, education, occupation, and other characteristics. Even though the sample has an average saving-income ratio much higher than that for all households, the difference in saving nearly disappears when differences in income are taken into account.² Various characteristics of the sample probably magnify the effects of pension coverage, whatever those effects may be. The higher-income and better-educated groups are likely to be aware of their pension coverage, what it costs and what benefits it bestows. Other groups are likely to be hazy about these costs and benefits and less inclined to alter their saving behavior when they become covered. This sample may give us a glimpse of reactions to pension coverage in the years ahead, as the persons already covered and those to be covered gradually become fully aware of the costs and benefits of pension plans. Although this is far from a random sample, therefore, it need not be dismissed as irrelevant to the population at large; indeed, there is some justification for believing it parallels the U.S. population of the future.³

Finally, does the time period covered introduce any bias? During the twelve months preceding the spring of 1959, to which the second questionnaire and most of our data pertain, the economy recovered from a recession; during the first part of that period, it had less than full employment. This recession had a mild effect on personal saving in the economy at large, judging by 1958 in comparison with 1957. Yet a small change in propensities to save could affect the comparison of covered and not-covered households—if it affected them differently. We do not know whether it did or not. It might be conjectured that the

² The average saving-income ratio in the sample, excluding pension contributions entirely, is 13.7 per cent. (This average excludes incomplete questionnaires and the self- and not-employed, but includes households with large gains or losses and extreme saving ratios excluded later.)

The corresponding percentage estimated for sample households in the 25-49 age group (\$5,000-7,499 income class) and not covered by pension plans is 6.7, about equal to the U.S. average of 6.2 for 1958. (The sample for this estimate excludes households with large gains or losses and extreme saving ratios, in addition to those self- or not-employed. These exclusions make the estimate lower than it would otherwise be.)

³ For further comparisons of the larger population of Consumers Union members, from which the present sample came, with all households, see F. Thomas Juster, "The Predictive Value of Consumers Union Spending-Incentive Data," in *The Quality and Economic Significance of Anticipations Data*, Princeton University Press for NBER, 1960, esp. pp. 274-279.

recession, while it lasted, induced workers (wage earners more than salaried workers) employed in cyclically sensitive sectors to increase their saving in preparation for possible layoffs. This increase would affect our results only if the fraction of sample households employed in cyclical sectors was not the same for the covered and not-covered groups.

The covered group in the sample is heavily weighted with teachers and government workers, who work in a relatively noncyclical sector. These occupational groups aside, however, most covered households in manufacturing and transportation might be concentrated in industries hit hard by recessions. If so, during the survey period of recession, covered workers would save more as a group than they do in normal years, and so would appear to save more relative to workers not covered. This effect can be checked by looking at salaried workers and wage earners separately, since the latter should respond to recessions much more strongly than the former, and by comparing the difference between covered and not-covered workers in these groups with the difference for teachers and government workers, where the effect of recessions should be weak. Most of these limitations of the sample, therefore, can be checked by appropriate stratification, and none necessarily detracts from the significance of these data as reflections of the present and future effects of pension plans.

2. Definition of the Saving-Income Ratio

The analysis which follows compares the saving of households in various subgroups to determine the effect of pension coverage on personal saving. Saving represents the change in net worth and is measured by the change in:

1. Checking and savings accounts, government bonds
2. Common stock and mutual funds (ignoring price movements)
3. Other marketable securities
4. Equity in annuities and life insurance (cash surrender value)
5. Equity in real estate (including own house or apartment)
6. Equity in pension plan (excluding social security)

less the change in:

7. Total debt (incurred for purposes other than buying a house)

The respondents indicated the change in each category by checking boxes marked "no change" and an increase or decrease of under \$250, \$250-499, \$500-1,000, and over \$1,000. If the change was over \$1,000, they were asked to write in the amount. Saving was computed by adding (or subtracting if a decrease) the midpoint of each box checked (reversing the sign of total debt) or the amount written in.

The analysis distinguishes between discretionary saving, pension saving, and other contractual saving. Total personal saving is the sum of all three, and total *other* saving is the sum of the first and third. Discretionary saving goes into cash balances and securities or pays off nonmortgage debt (items 1, 2, 3, and 7).⁴ Pension saving finances group plans (item 6). Other contractual saving (i.e., other than pension saving) finances insurance and annuities or pays off mortgages on real estate. These are standard definitions. Of course, any contractual payment is undertaken initially at the discretion of the payer, but most households no doubt feel that these payments, once started, have a budget priority over any desires to save regularly through additions to cash or securities.

These items exclude saving in the form of hand-to-hand currency (assumed negligible in this sample), social security taxes and employers' contributions to pension plans (discussed in Chapter 4), and consumer durable goods. (The self-employed, who may invest in their own business, are excluded, as is explained later.) A strong argument can be made for including purchases of consumer durables in saving, especially here, since we take account of debt, which is often incurred in making these purchases, though they are not likely to be as close a substitute for pension contributions as are most financial assets.⁵

Saving, as computed, includes assets purchased with funds other than reported income. Some households benefit from capital gains, gifts, or inheritances, and others suffer capital losses from stock declines or sick-

⁴ Increases in nonmortgage debt reduce discretionary saving. Although instalment debt payments, once undertaken, are more contractual than discretionary, they were combined with the latter on the ground that, since many households accumulate cash to make lump-sum purchases of durables, these accumulations should be offset by the debt payments of households that buy durables with instalment credit. The debt item, anyhow, is too small on the average to affect our results materially.

⁵ Chapter 3, Section 5, compares the results with computations that include these purchases.

ness and accidents. Such occurrences, when large, materially alter the normal financial condition of households and may be expected to disrupt their saving pattern appreciably. The inclusion of these households need not bias the analysis, but it might if by accident they were heavily concentrated in either the covered or the not-covered group. At any rate, it was deemed desirable to exclude them. Therefore, 938 households that reported unusual capital gains or losses over \$1,000 during the year were excluded. A few households that failed to answer the question (9B) correctly may have been missed, but this exclusion probably catches most of them.

Since the amount of saving by a household depends significantly on its income, saving may be expressed as a fraction of each household's income and the analysis may be confined to a comparison of these fractions for covered and not-covered households. Income represents the reported amount for the year 1958. Saving, as has been noted, covers the twelve months preceding the spring of 1959, so that ratios of saving to income computed from these data are probably higher than the true ratios for this period, because in the nonoverlapping first quarter of 1959 the economy was recovering from a recession and incomes were higher than in the first quarter of 1958. This upward bias in the ratios is probably small, however, and is not likely to affect the analysis. On the other hand, the income figure is the amount received before taxes, which is larger and produces a lower ratio than disposable income—a more appropriate concept.⁶ Reinforcing this downward bias in the ratios is the tendency for respondents to understate their saving, because in reporting saving over the past year they may forget about some of it. This is less likely to be true of reported income. None of these sources of bias, however, seems important for our purposes.

Taking the ratio of saving to income does not adjust entirely for income level, which appears to affect not only the amount of saving but its ratio to income as well. We can partly check for this by grouping the sample by income levels. Such an adjustment fails in so far as

⁶ Possible bias from using pretax instead of disposable income appears negligible (discussed in Chap. 3, Sec. 5).

A bias in the opposite direction results from the exclusion from income of the value of fringe benefits paid by the employer (including social security taxes), which might appropriately be considered a part of total income even though not of disposable income.

reported income differs significantly from "permanent" income: more of an increase in income may be added to saving than to consumption if the increase is considered temporary than if it is considered permanent. It is difficult to adjust for this effect, because it is necessary to know, in addition to actual income, the long-run level expected by the household at the current stage of its life cycle and how its saving adjusts to deviations of the actual from this long-run level. This problem will receive further discussion later; for the moment it may be noted that this difficulty is not likely to affect the analysis materially, since deviations of current from permanent income seem unrelated to households' pension coverage and so are distributed among sample households at random.

3. Households Excluded

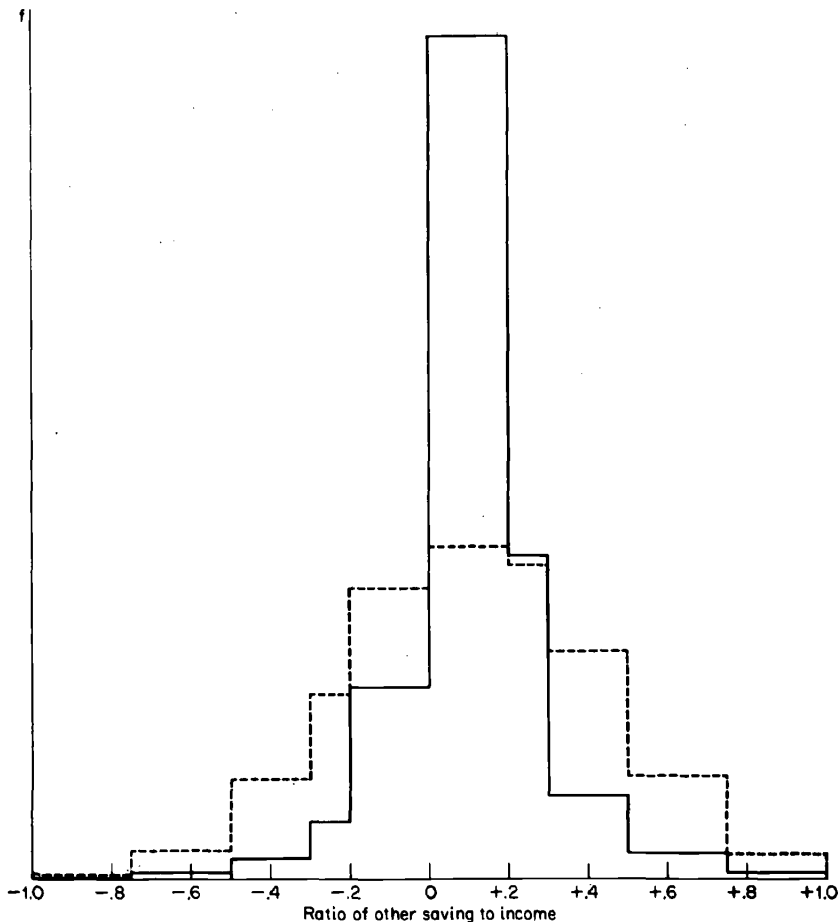
Of the 15,873 questionnaires returned, only 11,513 were used (see Appendix Table A). First, 938 households with unusual gains or losses over \$1,000 were excluded. Second, in 3,273 cases, the reported status of the head of household was self-employed, not employed (retired, housewife, student, or unemployed), or not given. These had to be excluded because none is covered by group pension plans (except possibly the unemployed) and most are not likely to save comparably to covered households. The self-employed are likely to have higher and the not-employed lower saving ratios than employed households that are not covered. Third, 149 additional questionnaires were incompletely filled out and had to be excluded.

The frequency distribution of the saving ratios for these 11,513 is shown in Chart I along with the normal curve having the same mean and variance (and covering the same intervals, which are not all the same size). The distribution of ratios is more peaked than the normal curve and also, what the chart does not show, denser at the ends. The distribution differs significantly (at the .01 level) from the normal curve, and does so even when the peak is suppressed by combining the ratios from -20 to +20 per cent into one interval. The important differences from the normal curve occur in the tails.

Partly because of this, and partly because extreme ratios in nearly all cases obviously reflect either unusual financial circumstances or

CHART 1

Frequency Distribution of the Ratio of Total Other Saving to Income for Sample Households (Solid) and Normal Curve with Same Mean and Variance (Dotted)



Note: Excludes contributions to group pension plans and social security taxes. Ratios over 1.0 and below -1.0 included in estimate of mean and variance, but not plotted.

gross mistakes,⁷ extreme ratios were excluded. The cutoff point was put at 50 per cent (excluding pension saving), thereby excluding 575 questionnaires having ratios of that percentage or more in absolute value. (Of the remaining 10,948 households, 8,027, or 73 per cent, were covered by a group pension plan.) The income distribution of these excluded households (and also those with unusual gains or losses, previously excluded) in Table 3 shows that they come from all levels and

TABLE 3
Distribution by Income of Households with Extreme Saving Ratios

Income	Ratio of Other ^a Saving to Income					
	Minus 50% and Under		Plus 50% and Over		Total	
	No.	Per Cent of Total	No.	Per Cent of Total	No.	Per Cent of Total
Less than \$3,000	10	6.8	9	1.3	19	2.3
3,000 - 3,999	10	6.8	18	2.7	28	3.4
4,000 - 4,999	15	10.3	32	4.7	47	5.7
5,000 - 7,499	48	32.9	160	23.7	208	25.3
7,500 - 9,999	27	18.5	156	23.1	183	22.3
10,000 - 14,999	23	15.8	202	29.9	225	27.4
15,000 - 24,999	13	8.9	77	11.4	90	10.9
25,000 and over	0	0.0	22	3.3	22	2.7
Total	146	100.0	676	100.0	822	100.0

Note: Includes households with unusual gains or losses over \$1,000, but excludes self- and not-employed.

^a That is, excludes contributions to group pension plans and social security taxes.

so do not reflect circumstances characteristic of households with high or low incomes, though they are, as is to be expected, somewhat concentrated in the extremes. The distribution of the total by income in the last column of the table differs from that for the reduced sample (Table 2) significantly at the .05 level, but only because of the comparatively high proportion of ratios minus 50 per cent or under in the

⁷ Such as writing in the amount of certain assets held rather than the change in the amount over the past year, as was asked for. Many of the excluded questionnaires may have this defect. The extreme ratios do not result from understatement of income used in the denominator, because the amount of income was written in, most likely with little error. It is more probable that most erroneous extreme ratios resulted from gross mistakes in the written-in amount of changes in assets, mainly through misunderstanding.

less-than-\$3,000 income class. If this class is excluded, the two distributions do not differ significantly.

Excluding the extreme ratios does not remove all households with unusual circumstances or all questionnaires with gross mistakes in reported saving, but probably eliminates most of them. Defining extreme ratios as plus or minus 20 per cent did not alter the subsequent results materially. Some support for selecting 50 per cent as the cutoff point is provided by the conformity of the remaining sample to the normal curve. The distribution of the remainder, with the peak suppressed as before by combining the ratios from -20 to +20 per cent into one interval, does not differ significantly from the corresponding normal curve even at the .10 level. In view of this, we may test for the significance of differences in the subsequent analysis on the supposition that the sample was drawn from a normal population.

These exclusions from the original sample all tend to reduce the amount of variability in the reduced sample and thereby to increase the significance of any observed differences. Yet the reasons for each exclusion seem legitimate for our purposes, so that the results do not seem to be biased and indeed should be more reliable. For we are interested in differences between the normal saving patterns of groups that differ only in their pension coverage, and unusual financial circumstances or habits, whatever the cause, only hide the basic differences we are looking for.