

NBER WORKING PAPER SERIES

WHAT PEOPLE DON'T KNOW
ABOUT THEIR PENSIONS
AND SOCIAL SECURITY:
AN ANALYSIS USING LINKED DATA
FROM THE HEALTH AND
RETIREMENT STUDY

Alan L. Gustman
Thomas L. Steinmeier

Working Paper 7368
<http://www.nber.org/papers/w7368>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1999

This paper is funded by a grant from the National Institute on Aging to the National Bureau of Economic Research. It is part of the NBER programs in Labor Studies and Aging. We would like to thank Olivia Mitchell for her helpful comments. Any opinions expressed are those of the authors, and not of the National Institute on Aging or of the National Bureau of Economic Research.

© 1999 by Alan L. Gustman and Thomas L. Steinmeier. All rights reserved. Short sections of text, not to

exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What People Don't Know About Their Pensions
and Social Security: An Analysis Using Linked
Data From The Health and Retirement Study
Alan L. Gustman and Thomas L. Steinmeier
NBER Working Paper No. 7368
September 1999
JEL No. D31, E21, H55, J14, J26, J32

ABSTRACT

Pension plan descriptions from respondents to the 1992 Health and Retirement Study are compared with descriptions obtained from their employers. Earnings histories reported by respondents are compared with earnings histories from the Social Security Administration. The probability of linking employer pension data, which is two thirds for current jobs, and of obtaining permission to link an earnings history, which is over 70 percent, are not well explained by respondent characteristics.

Half of respondents with linked pension data correctly identify plan type, and fewer than half identify, within one year, dates of eligibility for early and normal retirement benefits. Benefit reduction rates are essentially not reported. Respondents do better in reporting pension values, but the unexplained variation is still considerable. In contrast, respondent reported values, together with other observables, account for 80 percent of the variation in pension values and 75 percent of the variation in covered earnings measured from linked records. Thus prospects are good for imputing plan values, but not for imputing the location or size of early retirement incentives. Our findings raise questions about how well respondents understand complex pension and Social Security rules.

Alan L. Gustman
Department of Economics
Dartmouth College
Hanover, N.H. 03755
and NBER
alan.l.gustman@dartmouth.edu

Thomas L. Steinmeier
Department of Economics
Texas Tech University
Lubbock, TX 79409
wbtlsttacs1.ttu.edu

I. Introduction

It is now widely recognized that analyses of retirement and saving behavior must take account of pensions and Social Security. For those approaching retirement age, on average their pensions and Social Security are as valuable as the total of all their other assets combined (Gustman, Mitchell, Samwick and Steinmeier, 1999). Moreover, both pensions and Social Security significantly affect the rewards for continued work.¹

It is not surprising then that surveys are including more questions about pensions and Social Security. Some surveys, especially those of people approaching retirement age, not only ask respondents about their pensions, but also collect pension plan descriptions from the respondents' employers.² To improve the quality of information on these key data, the Health and Retirement Study links the respondent survey to Social Security earnings histories for the three fourths of respondents who gave permission allowing the Social Security Administration to provide the records.

In this paper we discuss similarities and differences in the information obtained from respondent surveys and from linked data. First, we compare and contrast pension plan characteristics and values when they are computed from respondent versus employer reports. Second, we compare and contrast Social Security benefits estimated from the respondent's self reported work history with benefits estimated from the respondent's actual covered earnings

¹For relevant literature reviews, see Gustman and Juster (1996), Lumsdaine (1996) and Lumsdaine and Mitchell (forthcoming).

²Surveys that link pension plan descriptions obtained from employers to respondent provided data include the Survey of Consumer Finances for 1983 and 1989, the National Longitudinal Survey of Mature Women, The Health and Retirement Study, and soon the Panel Study on Income Dynamics.

history.

Using the comparison of respondent, firm-provided and administrative information, we determine how well respondents understand their pensions and how accurately they report their covered work history. Further, we devise methods to predict pension and Social Security benefits for respondents when appropriate data are not available. More specifically, we address the following key questions:

- How do the covered, employed respondents for whom there is a matched employer-provided pension plan description, and the respondents for whom there is a matched Social Security record, differ from those with missing pension or Social Security records?
- What do comparisons between benefit amounts self reported by respondents with and without matched records suggest about the biases from confining samples only to those for whom matched records are available? For pensions? For Social Security?
- Earlier studies suggest that respondents frequently misreport pension plan type (Mitchell, 1988; Gustman and Steinmeier, 1989). What problems are created by such misidentification when linked data are missing and the researcher must rely exclusively on respondent reports?
- Among those for whom we have both respondent-reported and linked employer pension data, how do the plan features reported by respondents compare with those computed from employer provided pension plan descriptions? How do plan values compare? Are studies that rely exclusively on respondent reports of these key plan features or plan values subject to serious bias?
- What are the effects on the sample if analysts only include in their studies observations

for which matching employer provided and Social Security data are available, and omit observations for which plan descriptions are only available from respondents?³

- How do Social Security benefits differ when estimated from respondent reports of earnings histories from estimates based on administrative records of covered earnings?⁴
- What do the comparisons between benefit amounts computed from respondent-reported and matched data suggest is the best way to impute pension and Social Security benefits for those respondents without a linked pension or Social Security record?
- Do respondents do better in describing their plans when they are closer to retirement than when they are further away?

Once these questions have been answered, we turn to another aim of this paper, to facilitate research with the linked employer provided pension plan descriptions and the Social Security records. Some researchers may feel they must either use only those observations that have linked data available, or estimate relations for the full sample using only respondent reports.⁵ If only observations with linked data are used, the sample may not be representative. But if respondent data are used, errors and imprecisions may be introduced, especially in analyzing retirement and saving behavior. We report equations that relate pension and Social Security values estimated using linked employer provided pension or Social Security records, to

³For example, in their reduced form retirement study based on the HRS, Gruber and Coile (1999) eliminate observations when linked pension or Social Security data are not available.

⁴Bernheim (1988) found, using a different data set, that modal predictions of Social Security benefits from self reported data were unbiased. The findings clearly suggested, however, that there were important differences throughout the distributions.

⁵There is an additional problem because many respondent reports are incomplete, leaving out details of pensions and earnings history.

the values obtained from respondent reports and other independent variables. If we take the outcome produced using linked data as a more accurate measure, these equations allow researchers to project pension outcomes for those observations for which linked data are not available. In addition, we will produce files for distribution to HRS users that contain the predicted values of these variables for all observations, including the observations for which linked data are available. This will make information from the linked data available to a wider group of researchers, including those who do not apply for or obtain permissions to use the linked data directly.⁶

There also is a broader set of questions about behavior. To the extent that people do not understand their pensions and Social Security, there may be consequences for measurement of wealth and retirement incentives, and also for how behavioral models should be specified. If older workers have a poor understanding of pensions and Social Security incentives, then those who are misinformed may not respond to measured incentives in the way that models of saving and retirement behavior would predict. That is, if respondent reports differ from employer-provided plan descriptions, it is not clear whether the researcher should use linked data in estimating behavioral models on the assumption that all actors are well informed. In subsequent research we will determine the extent to which any misinformation influences behavior. That, in turn, requires

⁶A procedure has been established at HRS, in coordination with the National Institute on Aging and the Social Security Administration, to protect respondent confidentiality when linked data are made available. A researcher must apply for access to restricted data. The application requires a research plan, a data protection plan, a demonstration of grant support from a federal agency, and a promise not to link the restricted data to any but specified respondent files. The basic respondent survey without linked data is available on an unrestricted basis. The pension and social security data file produced from this work will provide information derived from the linked data, but will be made available on an unrestricted basis.

that we determine which features of these plans cause the most confusion, carefully detailing the extent of misunderstandings. We intend to determine which features cause the most confusion in this paper.

II. The Data

The Health and Retirement Study (HRS) is a longitudinal, nationally representative study of older Americans. The survey began with an initial cohort of 12,652 individuals from 7,607 households, with at least one household member born from 1931 to 1941. The first wave was fielded in 1992.

In wave 1, 71.6 percent of respondents gave permission to link Social Security earnings histories to their interview record. Records were actually linked for 95 percent of those who gave permissions. The share of respondents with linked Social Security records was raised to 74.9 percent as a result of additional permissions obtained in waves 2 and 3.

Social Security values may be estimated from respondent reports of their earnings and work history. Specifically, respondents were asked about the starting date on their current job, starting and ending dates for their last jobs, (i.e., the job last held by those not working in 1992), and starting and ending dates for the previous 5 year job held before the current or last job. Respondents were also asked about earnings at these dates. In addition, the survey asked respondents in Wave 3 about the date of entry into the labor force, how many years were worked before the date the previous job was secured, and how many years of work were in jobs covered

by Social Security. From this information, we construct a covered earnings history⁷ and we use it as the basis for calculating the respondent's Average Indexed Monthly Earnings (AIME), and the Social Security benefit the respondent is entitled to (Primary Insurance Amount, or PIA).

If a respondent was covered by a pension, the HRS also requested from the employer a detailed description of the pension plan. Employer-provided pension plan descriptions proved to be more readily available for a respondent's current job than for jobs held in the past. The match rate was 65% for the 4456 jobs that respondents held at the time of the survey, where they reported they were offered pensions. The match rate was 66% for the 1387 cases for the last job held by respondents with no current job in 1992, who reported a pension on their last job. For the previous job of 5 years duration held before the current or last job, the match rate was 35% of 2839 pensions. In addition, the survey identified another 750 jobs with pensions. Since the survey did not ask for the employer name and address for those jobs, none of the jobs had matched pensions.⁸

Averaging over all these jobs yields an exact match percentage for employer-provided plan descriptions of 51%. This combines a figure of almost two-thirds from current and last job pensions, with a much lower figure for pensions on previous jobs.

⁷Wage profiles are forced through all observations, and values for missing years are projected off the profiles on the basis of experience and education. The wage profile coefficients are taken from Anderson, Gustman and Steinmeier (1999) and are based on data from the Survey of Consumer Finances. Coefficients are: experience .0138221, experience squared -.0002827, and experience * education .000996.

⁸These counts are for individual jobs, not individual pensions. That is, if an individual had both a DB and a DC pension in a job, it would be counted once in these tallies. However, an individual may have more than one entry if he or she had pensions in both the current job and a previous job, or in the last job and a previous job.

Respondents were also asked a series of detailed questions about their pensions. First the respondent was asked about pension coverage, and if covered, about plan type. If the respondent indicated his plan is of the defined benefit variety, the question sequence asked about the dates of early and normal retirement, the year of expected retirement, and the associated yearly benefits. If the respondent indicated his plan is of the defined contribution type, the questions focused on the amounts currently in the account and on rates of contribution.

In the case of those with defined benefit plans, the Summary Plan Description (SPD) provided by the respondent's employer contains a full and accurate representation of the pension. To be sure, there are some possible sources of error. Most importantly, some plan descriptions will have been dated, referring to provisions in place some time before 1992, but no longer relevant in 1992. Also, despite extensive checking, some plan features may have been miscoded. Moreover, if a firm has experienced complex merger activity, multiple plans may cover different individuals with similar apparent backgrounds, creating the possibility of a mismatch.⁹

In the case of defined contribution plans, the respondent report as to the balance in the account may be more accurate for estimating plan value. Because the HRS must preserve respondent privacy, when collecting pension plan descriptions, the name of the covered individual was not identified to the firm. Therefore, the amount accumulated in the pension account held by a particular respondent must be estimated from the firm's contribution rate as reported in the plan

⁹The pension plans were coded from the Summary Plan Descriptions provided by employers at the University of Michigan, and software written for that purpose was used to evaluate the coded plan descriptions. The software generates pension values by applying the coded rules from the pension plan to respondent reports of earnings and tenure. The user specifies assumptions as to interest rate, wage growth, and which respondent reports of retirement dates are to be used. Details about the procedures employed are available in Gustman, Mitchell, Samwick and Steinmeier (forthcoming).

description, and the respondent's self reported work history and contribution rate. In addition, although the HRS collects the plan description at a moment in time, both the contribution rates and the returns may have changed over time. Nevertheless, the respondent's answer may be subject to reporting error, perhaps to substantial error in some cases.

III. Empirical Analysis

A. Matching Rates for Pension and Social Security Records.

Table 1 reports the results of a probit, where the dependent variable is a qualitative indicator equal to 1 if permission was given by the respondent in wave 1 allowing HRS to match the Social Security record. Because the Social Security Administration was able to match 95 percent of the records of those who gave permission, to understand what determines the match rate, one needs to understand what determines the permission rate. The independent variables are various characteristics of the respondent and the respondent's job.¹⁰ The partial effects of the probit are reported together with the z statistics. The independent variables account for only a small part of the variation in the Social Security permission rate.¹¹ Among the individual variables that are significant, black and Hispanic respondents, respondents with higher assets and highest education, respondents who expect never to retire, or who do not report a specific retirement date, were less likely to grant permission.¹²

¹⁰Appendix Table 1, column 1, reports the Social Security permission rates in wave 1 for respondents according to individual and job characteristics.

¹¹Haider and Solon (1999) and Olson (forthcoming) reach similar conclusions.

¹²Despite the relatively poor fit in the Social Security permission equations, we will include these covariates in later regressions to adjust predictions for whatever selection is systematically associated with these observables.

Table 2 reports probits where the dependent variable is an indicator of whether there is an employer-provided pension plan description for the respondent.¹³ The first column indicates the correlates of pension matches for all respondents who indicated they were covered by a pension on some job. Separate equations are shown for pension matches on the current job held in 1992, the last job for those with no current job, and the last five year job held previous to the last or current job. For the purpose of examining retirement incentives, it is the pension on the job just preceding retirement that is most important; though the survey has a lower pension match for previous jobs, this is less crucial for retirement modeling.

The likelihood of finding an employer plan description is not closely related to the independent variables included in Table 2. However, the fit is somewhat better than it was for a Social Security permission rate in Table 1. Across all pension plans, blacks, those with more schooling, homeowners, those with the shortest planning horizons, those with the longest tenure, and those with jobs in nonmanufacturing, are more likely to have a pension match. Those with the highest assets and earnings, those from firms employing fewer than 100, those in management jobs, and those who report they are covered by DC plans, are less likely to have a matched plan description.

In addition, the more valuable the plan, the higher the probability of a match. This last finding suggests that if one confined the sample only to those with matched pension plan descriptions, the value of the pension will be overstated. Similarly, because the probability of a Social Security earnings record match is lower the higher an individual's earnings, if a researcher

¹³Appendix Table 1, column 3, reports the pension matching rates in wave 1 for respondents according to individual and job characteristics.

confines the sample to those with a Social Security match, he is omitting from the sample a disproportionate share of those who will receive the highest Social Security benefits.

B. Pension Comparisons

Table 3 contains four tables describing the joint distribution of plan types reported by respondents and their firms on jobs held by respondents in 1992. These data are only for respondents with a matched employer-provided pension plan description.

Table 3A contains the frequencies of plan type reported by covered respondents and by their employers. If the firm reports two plans, one a DB and one a DC, the firm is classified as offering both, and there is one entry in the “Both” row. In Table 3A, we find 1881 cases where both the individual and the firm reported having a DB pension, with or without a DC pension: $(777+380+409+315)$. For DC plans the comparable number is 916, where both the respondent and employer reported that the individual is covered by a DC plan $(327+111+163+315)$.

The share of the total of 2907 observations falling in each cell is reported in Table 3B. The observations along the diagonal in Table 3B represent only about half the joint distribution of firm and individual reported plan types, suggesting that respondents do a poor job of reporting the type of pension plan they are covered by. Some of the misreporting may be due to nonparticipation; even when the firm reports it offers both plan types, some 14 percent of respondents report they are covered only by a DB plan. This difference may reflect a failure of the respondent to report coverage by a DC plan when they don’t participate in the DC plan. Even allowing for this misreporting, the discrepancy in plan types is substantial, amounting to over one third of the plans. Because the sequence of questions asked about the pension is conditioned on the plan type reported by the respondent, this is a crucial problem for surveys that wish to

determine pension values solely from respondent reports.¹⁴

Table 3C reports the distribution of respondent reports of plan type, conditional on the plan type reported by the firm. As seen in the first column of Table 3C, among respondents whose employers report they are covered by a defined benefit plan, only 56 percent report that their plan is DB. An additional 27 percent of those whose employers report their plan is DB report that their plan is both DB and DC. Thus among those whose employers indicate they are covered by a DB plan only, 83 percent of the respondents report coverage by a DB plan. Fifteen percent of those whose employers report they are covered by a DB plan report they are covered by a DC plan only. From the second row of Table 3C, among those whose employers indicate they are only covered by a DC plan, slightly more than half, 54 percent, get the plan type right. Altogether, among the observations where the employer reports a DC plan only, the respondent reports a DC plan, either held exclusively or (mistakenly) together with a DB plan, in 72 percent of the cases. Twenty six percent of those whose employer reports their plan is DC, representing 14 percent of the sample with a matched pension plan, report their plan is DB only. For those whose employers report they are covered by both a DB and a DC plan, slightly more than a third, 35 percent, report coverage by both a DB and a DC plan. Among the remainder, 45 percent

¹⁴Means from the separate distributions suggest more agreement between respondent and firm reports of plan type than is found in the micro data. At the micro level respondents and firms agree on plan type in 49 percent of the respondent-employer matches. In contrast, compare the fractions with only DB, only DC, or both types of plans computed separately from the reports of individuals and firms. From the bottom row in Table 3B, we see that 46 percent of respondents report that they are covered by DB plans only, 24 percent report they are covered by DC plans only, and 28 percent report they are covered by both types of plans. Similarly, from the last column of Table 3B, 48 percent of firm provider reports indicate that the plan is DB only, 21 percent indicate that the plans are DC only, and 31 percent indicate that the firm offers both types of plans. Thus when the descriptive statistics from firm and individual reports are compared, they suggest a much higher level of agreement than is found at the level of the individual observation.

report they are covered by a DB plan only, and 18 percent report they are covered by a DC plan only.

By way of comparison, we report the percentages from Gustman and Steinmeier (1989, Table 6), which provided results analogous to Table 3C using the 1983 Survey of Consumer Finances (SCF). In the SCF, 63 percent of respondents whose firms reported a DB plan only also reported coverage by a DB plan only, compared to 56 percent from the HRS in Table 3C. In the SCF, of respondents whose employers reported a DC plan only, 37 percent reported they had a DC plan only. This compares to 54 percent in Table 3C. Although the HRS respondents do a bit worse in identifying DB plans only, they do much better in identifying DC plans only than respondents in the SCF sample did, perhaps due to the increasing popularity of DC plans over the decade.

Suppose for those without a matched pension we wish to determine plan type. The simplest idea would be to use the respondent report. But the respondent report will not be reliable for two reasons. First, there is a problem of selection bias; that is, the sample without a matched pension is different from the sample with a matched pension. Second, the respondent's report will be wrong for roughly half the cases. Consider the distribution of firm reports conditional on respondent reports as shown in Table 3D. For those respondents who report they have a DB plan only, 58 percent of the firms indicate they offer only a DB plan. Counting those whose firms report DB and DC coverage, DB plan descriptions will be available for 88 percent of those respondents reporting DB coverage, but DC plans will be missed for the additional 42 percent of those whose firms report they have only DC plans or both types of plans. Among respondents who report they have only a DC plan, 47 percent will have an exact match provided

by an employer who reports offering a DC plan only. But we will have DB plan descriptions from employers for 53 percent of respondents who report coverage by a DC plan only. Among respondents reporting coverage by both types of plans, descriptions of both plans are available from only 39 percent of their employers, suggesting about 60 percent of respondents are misinformed.

Ages of Eligibility for Early and Normal Retirement Derived From Self Reports and Firm Reports

Ages of early retirement are reported in Table 4 for respondents in jobs held in 1992, where both the respondent and the firm report the plan is defined benefit. The median age of early retirement reported by the respondent is 57. When applying the formula reported by the firm to respondents' self-reported work histories, the median age of early retirement across plans is 55. The average early retirement age is 57.6 when reported by the respondent, and 55.4 when calculated from the firm-provided plan description. That is, when based on the firm-provided plan description, the early retirement date is about two years earlier than the early retirement date reported by the respondent. According to the firm provided data, two thirds of respondents will be able to retire early by the time they reach age 55; but less than half of respondents think they will be able to retire early by age 55. Thus respondents to the HRS appear to be more pessimistic about their eligibility for early retirement than is warranted by the provisions of their plans. We will see below that some respondents may simply be ignorant of the opportunity to retire early, or perhaps otherwise consider it to be irrelevant. SCF respondents ten years earlier did not overstate the age of early retirement. In the data from the 1983 SCF, we found that the median early dates expected by respondents and the median early retirement dates computed from plan provisions

and earnings histories were both age 55. Moreover, the SCF included workers from age 40 on, so that those sampled in the SCF were further from retiring than the HRS sample. Yet the median expectations of early retirement date in the SCF sample were more accurate than in the HRS sample. One possible reason is that early retirement was somewhat less common in the SCF than in the HRS, and the age of early retirement was not as far below the age of normal retirement as it is in 1992. Thus even if a number of respondents were ignorant of the opportunity to obtain early retirement benefits in the SCF, it made less difference.

Another piece of evidence suggests that people are less optimistic today about when they can retire than they were in the past. In previous work using SCF data, we found that in the self-reported data, mean early retirement dates were lower than the median dates by about 3 years. Because the medians were equal between the self reported and firm reported data, this suggested that there were a few people who were highly optimistic about when they would be eligible for their early retirement benefits. In the HRS, evidence of this optimism has disappeared: mean and median dates of expected eligibility for early retirement benefits are the same. This result corresponds with the findings from the direct comparisons of early retirement dates based on the respondent and firm-provided data. People are more pessimistic about the ages of eligibility for early retirement benefits in the HRS than they were in the past.

It is also useful to examine the distribution of differences between the early retirement dates expected by respondents and the dates they will attain early retirement eligibility according to their firm reports. Of the observations with firm-reported retirement ages in the range from age 50 to 65, only 28 percent (435/1569) lie along the diagonal in Table 4, indicating that the respondent and firm based early retirement dates are identical. Moreover, only 43 percent of

respondents (671/1569) report an early retirement date within one year of the firm-based early retirement date. The simple correlation between the provider reported and self reported retirement dates is 0.353.

This wide variation in dates of eligibility for early retirement benefits is a particularly important problem for analysts wishing to model retirement. For many pension plans, benefit rules give covered workers a substantial benefit increment from working up to the point of early retirement age. By working the year a respondent becomes eligible for early retirement, the HRS respondent typically increases the present value of the pension by an amount equal to a year of pay, or more (Gustman and Steinmeier, 1999a). This provides a powerful incentive for respondents to remain at the firm. If, in formal models of retirement behavior, individuals appear to leave their firms before reaching that date, retirement models will indicate that individuals are not sensitive to economic incentives when making their retirement decisions. As seen in the responses tabulated in Table 4, many individuals will report an early retirement date that is later than the date they actually will become eligible for early retirement benefits. As they approach retirement age, in many cases the firm will make clear to them that they are being too pessimistic. But that new information may not be reflected in the individual responses to the survey, and the analyst will not have the correct answer unless the firm provided plan description is available. Unless an adjustment is made when using respondent reported data for those who seem to leave just before becoming eligible for early retirement benefits, this form of reporting error will cause the effects of pension incentives on retirement to be understated, and the parameter estimates will also lead to an understatement of the influence of Social Security on retirement.

How accurately are pension rules described by those respondents closest to retirement?

Table 5 compares the respondent-reported early retirement dates with the dates calculated from the firm plan descriptions and respondent reported records, this time confining the sample to the 371 respondents in the 1992 wave 1 HRS survey who indicated that they wanted to retire by 1995. For this sample, and using the pension formula reported by the firm to the respondent's self reported work history, the median age of early retirement eligibility is estimated to be 55, the same as the median age of early retirement eligibility expected by respondents. For the full sample, we found the median early retirement date expected by respondents to be two years later -- at age 57. Thus for those expecting to retire within three years of HRS wave 1, the firm-reported and respondent-reported median dates of early retirement correspond. However, a systematic bias persists. When the sample is confined to those who are approaching retirement, in the mean early retirement date based on respondent data is later than the mean date based on firm data: the average respondent reported early retirement age is 57.0; while the mean age of eligibility for early retirement calculated from firm reported data is 54.8. Moreover, the mean exceeds the median early retirement age in self reported data, suggesting that some of the respondents are overly pessimistic, but it is not those in the middle of the distribution. Thus individuals within three years of their expected retirement date also expect to have a later early retirement date than they in fact are eligible for, but unlike the full sample, the difference is due to overly pessimistic projections by some in the tail or tails.

When in Table 5 we confine the sample to those who intend to retire by 1995, we continue to observe the same wide discrepancy between the early retirement dates computed from provider formulas and from respondent data that characterized the full sample in Table 4. Thus for the full sample, 28 percent of the observations were found to project their early retirement dates

accurately, so that observations lie along the diagonal in Table 4. Similarly, for those who expect to retire by 1995, 27 of the observations lie along the diagonal in Table 5. In addition, in Table 5, we find that 39 percent of the observations for this group are within one year of the early retirement date calculated with the firm provided plan provisions, the same proportion found for all respondents. The correlation coefficient is 0.359 for the data in Table 5, virtually identical to the correlation coefficient of 0.353 found for Table 4.

Table 6 examines the joint distribution of *normal* retirement dates predicted from provider plan descriptions and from respondent data. The median normal retirement age is 62 in both the respondent provided data and the firm-provided plan descriptions. The means are 61.3 and 60.7 when normal retirement ages are computed from firm reported plan provisions and respondent expectations respectively. However, the discrepancies are much wider in the individual data than they appear from the medians or means. Among those with normal retirement dates between 50 and 65, as computed from employer-provided plan descriptions, 33 percent of the observations lie along the diagonal in Table 6. Among this same group, 40 percent have an expected normal retirement age within one year of the date calculated from employer-provided data. The correlation between the normal retirement dates from provider and self reports is 0.352. This suggests that while there is no systematic bias in the expectations about the age of qualifying for normal retirement, the dispersion is no lower for age of normal retirement than it is for the age of qualification for early retirement benefits.

A key characteristic of the defined benefit pension is the benefit reduction rate. It is the rate at which benefits are reduced for each year the covered worker retires before the normal retirement date. If the reduction rate is relatively low, early retirement is subsidized so that the

benefit accrual peaks at early retirement age.

The benefit reduction rate is available from both the individual and the firm for relatively few of the 1881 jobs where both the respondent and the firm report coverage by a defined benefit plan. As can be seen in Table 7, we can compute early retirement reductions for only about 58 percent of firm observations; most of the remainder are cases where the worker joined the firm recently enough that he or she is not eligible for early retirement prior to the normal retirement age. The median benefit reduction rate reported by firms is 4 to 5 percent. Respondents themselves have almost no idea how much the benefit reduction rate is. Out of 1881 respondents with a defined benefit pension in their 1992 job, 1512, or 80 percent, either do not think they are eligible for early retirement or do not know what the benefit reduction rate is for their plans. Indeed, it appears that half of the respondents do not think they can retire before the normal retirement age. Of the remaining 369 respondents, the median reported benefit reduction rate is 5 to 6 percent, well above the average observed from provider data. Altogether, there are only 234 observations on benefit reduction rates that are jointly reported by the respondent and by the firm. Of these, only 26 lie along the diagonal. For this limited number of observations, the correlation coefficient is 0.524.

As Table 8 shows, there is agreement in a majority of cases about whether a DC plan allows voluntary contributions.¹⁵ In 392 out of 873 observations, providers and respondents agree there are voluntary contributions. In another 121 cases, both respondents and the firm agree that there are no voluntary contributions. Thus, 59 percent of the cases lie along the main

¹⁵The self-reported voluntary contribution variable is a dummy variable for whether the plan has a 401(k) component. As such it would be incorrect to infer that a plan with voluntary contributions does not also have required contributions.

diagonal. With respondents and firms disagreeing about voluntary contributions in 41 percent of the cases, the most likely misreport occurs where the firm reports there are voluntary contributions and the respondent reports there are none. These misreports account for 35 percent of the observations.

Plan Values Derived From Self Reports and Firm Reports

Table 9 examines results for pensions from the current job at the time of the survey. The results for the first two columns are for individuals who said they had defined benefit plans whose value could be computed from the respondents' answers and whose firms indicated that the individual was indeed covered by a defined benefit plan. Combination plans with a defined benefit component are also included in these numbers. The last two columns pertain to individuals who said that they had defined contribution plans and whose firms indicated likewise. Individuals who reported they had both defined benefit and defined contribution parts to their pension, and whose firms agree that the pension had both components, may be in both columns. Individuals who say they had only defined benefit plans but whose firms indicate only defined contribution plans, or vice versa, are not included in this table, although they are included in later tables. Thus, this table and the next several tables pertain to individuals for whom the respondent and the firm agree on the type of pension, and for whom the principal ingredients of the pension (earnings, expected benefits, accumulations, etc.) are available and not imputed.

Table 9 shows the dollar amounts associated with various points on the univariate distributions of defined benefit and defined contribution pension amounts, calculated both using

the respondents' answers and the formulas in the pension documents obtained from the firms.¹⁶ For defined benefit plans, the survey asked when the respondent expected to start collecting the pension, and how much the pension would be. The pension value in column 1 of the table is the present value (discounted to 1992) of the expected benefits from the date of expected retirement forward, assuming that the pension remained the same in nominal terms.¹⁷ If the respondent failed to answer either the expected age when the pension would start or the amount of the pension, the pension value is not imputed. The second estimate of the pension value, which is in column 2 of the table, comes from applying the respondent's earnings and tenure at the expected collection date to the rules found in the pension documents.¹⁸ These rules give the value of the annual benefit, year by year. The value of the pension is again the discounted value of this stream of benefits.¹⁹ If the respondent's 1992 earnings were imputed, the value of the pension is not

¹⁶In Table 3 there were 1923 plans where both the respondent and the firm reported plan type as defined benefit. Table 9 reports pension values for 1122 defined benefit plans. The difference is accounted for by missing data, primarily because only two thirds of respondents in Table 3 reported expected pension amounts (or percentage of pay). For similar reasons, Table 9 reports results for 641 observations, whereas there were 916 observations in Table 3 where both the respondent and the firm reported there was a defined contribution plan.

¹⁷ By assuming that benefits stay constant in nominal terms, we ignore the fact that some pensions have benefits that are reduced when the participant becomes eligible for Social Security. About 12 percent of participants in defined benefit plans indicated that their pensions would be subject to such provisions.

¹⁸The projected future earnings are calculated using Social Security Administration projections, increasing the 1992 earnings by 6.3% per year.

¹⁹Some difference might arise because the respondent might assume future wage growth different than the 5% growth (the Social Security Administration's intermediate projection) we assumed when evaluating the firm pensions. Therefore, we adjusted the respondent's projected benefits to allow for implied differences in wage growth assumptions. To be more specific, the survey asked about the expected benefits on the date the individual expects to start them and the wages both in 1992 and at the normal retirement age. To project the wage to the benefit start

imputed, since the imputed earnings may not be a very accurate indication of actual earnings for individual respondents.

For defined contribution pensions, the value of the pension as determined by the respondent, which is in column 3 of the table, is simply the answer to the question regarding the current value of the accumulation. To arrive at the amount calculated from the plan documents, the required contribution amounts are calculated for each year that the participant has been employed in the firm. If this is expressed as a percent of the annual earnings, the percent is multiplied by the earnings in the year. These contributions are accumulated forward using a 6.3 percent nominal interest rate from the long run projections of the Social Security Administration, and the sum of these contributions is the implied present value. If the plan allows for voluntary contributions, the percentage of contributions in the survey year is extrapolated backward, and the individual is assumed to have contributed the same percentage of previous years' earnings. The plans themselves indicate if the contributions began after the respondent began to work for the firm. If an individual's 1992 earnings are imputed, or if the respondent did not respond to the

date, we interpolate the wage between 1992 and the normal retirement age. We then calculate the implied ratio of the benefit to the wage that the respondent is expecting at the benefit start date. Finally, we multiply this ratio by the wage the respondent would have if the 1992 wage grows by 5% per year. This essentially adjusts the reported benefit for differences in wage growth assumptions. However, there are some observations for which this procedure does not work very well. These observations have the common feature that the individual is almost at the normal retirement age but expects to work several more years. For instance, the individual may be 59, have a normal retirement age at 60, expect to retire at 65, and report earnings of \$40,000 at 59 and \$80,000 next year. Continuing this rate of growth until 65 yields a value which is clearly nonsense. Therefore, we compared the computed wage at expected retirement to the wage which would be obtained by simply growing the 1992 wage by the growth rate assumed in the Social Security projections. In the relatively few cases where the former exceeded the latter by more than 50%, we took the expected pension benefits at the expected retirement age at face value and did not try to make any adjustments.

question about contributions and the plan had voluntary contributions, the present value is not imputed. The resulting calculated defined contribution pension values are in column 4 of the table.

Table 9 indicates that for the defined benefit plans, the mean present value of benefits based on the employer reports, \$168,405, exceeds the mean value based on the respondent report, \$148,015. At the median, there is remarkable agreement between the defined benefit amounts based on the respondents' responses and those based on calculations from the pension documents. In fact, looking at the values in the 25th and 75th percentiles, the middle part of the distribution is almost identical. Only in the two tails are the pension amounts calculated from the firm documents noticeably higher than the amounts calculated from the respondents' expected pensions.²⁰ For defined contribution plans, however, the situation is considerably different. The mean pension value based on firm reports of \$85,790, greatly exceeds the mean value based on the respondent report, \$59,105. While the upper tail for the distribution of DB plans exhibits provider values that exceed the values reported by respondents, the same is true for the top 90 percent of DC recipients; that is, for DC pensions the *entire* distribution of accumulations that the respondents report is much lower than the amounts calculated from the plan documents. In the middle and upper parts of the distribution, the respondents' accumulations are less than half as much as the amounts from the plan documents at the corresponding percentiles.

²⁰In our earlier analysis of data for the Survey of Consumer Finances, we found that pensions based on respondent reports were more valuable than pensions based on the plan formula (1989, Table 5). The findings here are just the opposite. We also found that unlike our earlier findings, people on average are now more pessimistic about their ages of eligibility that is warranted by the plan. Perhaps the finding that people are less optimistic about their pensions than they were ten years ago reflects a shift of attitudes in response to changing firm policies over the past decade?

Although the distributions are similar for the defined benefit pension amounts calculated from the respondents' expectations and the pension documents, Table 10 indicates that there are substantial differences between the two amounts at an individual level. To some degree, this result could have been expected after looking at Tables 4, 5 and 6. If respondents are that inaccurate in reporting the early and normal retirement ages of their plans, it seems unlikely that they would be more accurate in reporting their expected pensions, which presumably involve more complex calculations than do the early and normal retirement ages. The fact that the joint distribution is more or less symmetrically distributed around the main diagonal reflects the fact that the two individual distributions are similar, and indeed the fact that the largest entries are down the main diagonal is encouraging. However, only about 40 percent of the observations fall along the main diagonal. Moreover, the ranges of the categories in this table are very wide, and even being one entry off the main diagonal is consistent with the respondents' estimate of pension value being half or twice the corresponding amount calculated from the pension documents.²¹

Table 11 shows the joint distribution for the defined contribution plans. The scatter of the plans is about the same as for the defined benefit plans, with 28 percent of the observations falling on the main diagonal.²² This is probably one of the main messages of both Table 10 and Table 11: at the individual level, there is a great deal of difference between the pension amounts based on the respondents' answers and those based on the plan documents. A second message is specific to defined contribution plans, namely, that the accumulations reported by respondents are systematically lower than the amounts calculated by using the pension documents to figure the

²¹The correlation coefficient, which is heavily influence by the larger values, is 0.615.

²²The correlation coefficient for the observations in Table 11 is 0.302.

contributions and cumulating these contributions over time. In this table this comes out in the fact that the majority of the entries are below the main diagonal. In Table 9 the difference in the mean value of the defined contributions plans amounted to almost fifty percent of the value reported by the respondents.

The systematic difference between the accumulations reported by the participants and the values calculated from the pension documents could have several potential explanations. Most obviously, the participants could be systematically under-reporting the plan accumulations. Alternatively, the rate of return used by the pension program in calculating the pension reports may be higher than the actual return experienced by the participants. Finally, for the plans with voluntary contributions, the participants may have contributed substantially less in prior years to the plan than their current contribution rates would suggest.²³ Note that the first of these explanations implies that the respondents' answers are systematically incorrect, while the last two imply that the amounts calculated by the pension program are incorrect.

To distinguish among these competing explanations, Table 12 presents the results for those DC plan participants whose plans do not allow for voluntary contributions. In Table 12, the number of entries above the main diagonal is approximately the same as the number below the diagonal. Although the number of observations is much lower than in Table 11, Table 12 suggests that among those participating in plans without voluntary contributions, there is no systematic tendency for the pension values calculated from the plan documents to be more or less than the accumulations reported by the respondents.

²³The pension calculation program assumes a constant contribution rate for participants of plans with voluntary contributions.

The fact that values appear close for plans without voluntary contributions, but not for those with voluntary contributions suggests that the third explanation mentioned above is correct. The other two explanations should have equal force whether the firm permits voluntary contributions or not. As for the accuracy of the reports, one would expect that participants in plans with voluntary contributions should be more accurate in their reports. For these participants, there is a periodic decision as to how much to contribute, and to the extent that this decision impels them to look at their pensions, they should have better information. As for the return assumptions built into the pension calculations, the fact that the calculated plan values and the reported accumulations agree for participants in plans with no voluntary contributions suggests that the return assumptions are approximately correct, at least for those participants. Those assumptions are for a very modest 2.3% real return. It is logically possible that the returns to those pensions with voluntary contributions are lower, but to produce the results in Table 11 the returns would have to be substantially negative for pensions with voluntary contributions.

This leaves us with varying contribution rates for plans with voluntary contributions as the primary suspect. In defined contribution plans without voluntary contributions, the plan itself specifies how much the firm and the participant contribute. This information is included in the pension plan calculations. The plan may also specify how these contributions have varied over time, and this information is also included in the calculations. For voluntary contributions, the plan may specify minimum and/or maximum contributions, but the actual contributions must come from the participant.²⁴ Further, if the contribution rate has increased over time for a participant,

²⁴ For reasons of confidentiality, the HRS did not tell the firm the name of the respondent when it collected the pension documents from the firm. It was thus unable to ask anything about the actual contribution rates for particular individuals and was forced to obtain this information

as might be expected as the participant moves closer to retirement, this information is not captured in the survey. Thus the calculations based on voluntary contributions for participants nearing retirement may considerably overestimate the values of the accumulations in the plans.²⁵ Note again that if this explanation is correct, it is likely that the accumulation reported by the respondent may be better than a plan value calculated from the pension document for these plans.

This point is expanded in Table 13 and in Figure 1. This table groups defined contribution pensions according to the value calculated by the pension program on the basis of the pension documents. The top part of the table deals with defined contribution pensions without a provision for voluntary contributions. The first row indicates the median accumulation reported by respondents, and the second row indicates the median value calculated by the pension program. The solid line in Figure 1 calculates the ratio between the two. The ratio is about 1.5 for small pensions (below \$5,000 in present value), and about 1.0 for pensions between \$5,000 and \$50,000. Above \$50,000, where the number of pensions is low, the ratio fluctuates more widely, but there is no evidence of a systematic bias. All in all, the solid line confirms our more casual finding from Table 11, that the reported accumulations and values calculated from the program do not differ systematically for pensions without voluntary contributions.

The lower part of Table 13, which is reflected in the dotted line in Figure 1, indicates the results for pensions with voluntary contributions. For these pensions, for values above the first category, there is evidence of a systematic difference between accumulations reported by respondents and those estimated by the pension program using the pension documents. There is

from the respondents.

²⁵The correlation coefficient for the data in Table 12 is 0.478.

some indication that the degree of underestimation is worse for larger pensions. As discussed before, the likely explanation for this is that the respondents have contributed less in previous years than they are currently contributing. This implies that for defined contribution plans with voluntary contributions, the calculated values from the pension documents probably overestimate the true value of the accumulations. This brings us to the question: what is the best estimate of the true value of the pensions? The inaccuracies in respondent reports are obvious. However, it is also likely that the values calculated from the pension documents are not completely accurate. Part of the reason is that these calculations still employ uncertain information from the respondents, such as earnings, years of service, and the like. Another reason is that the program is forced to make assumptions about information not collected from respondents, such as the time path of earnings and the time path of voluntary contributions. Is the best estimate of the true value of the pension the value calculated from the pension documents or the value calculated from the information that the respondents give? Or is it better to somehow combine the information in the two sets of numbers?

For defined benefit plans, the best estimate is almost certainly the one calculated from the pension documents. It is subject to errors in earnings and years of service, but it captures the details of the pension plan that most respondents may be unaware of unless they have already retired and asked the firm as to how the benefits were calculated. In cases where there are sharp differences, it seems much more likely that the respondent is making an uninformed guess than that the provider calculations are substantially wrong. The fact that the two distributions are similar does not give any cause to think that one or the other of the estimates is systematically biased.

Many of the same arguments hold for defined contribution plans with contributions that are completely specified by the plan. In addition to uncertainties regarding earnings and years of service, there is an additional uncertainty here having to do with investment returns. This uncertainty would be an argument in favor of the respondent reports of the accumulations. However, where there are large discrepancies between the accumulations as reported by the respondent and the firm provided plan description, again, there do not appear to be systematic differences between the two sets of values which would indicate that one or the other was biased.

For defined contribution plans with voluntary contributions, the situation is ~~even~~ murkier. It appears in this case that the amounts calculated from the plan documents are systematically higher than the accumulations reported by respondents. It further appears that the probable cause of this is that the respondents have increased their contribution rates over time, inducing the calculations from the plan documents to be too high. This would be an argument for giving more credence to the accumulations reported by respondents. However, there is still a tremendous amount of scatter in Table 11, which means that for a substantial number of respondents, there is an order of magnitude of difference between the amounts calculated from the pension document and the amounts that the respondents report. Although increasing contribution rates could contribute to some of the scatter, it seems unlikely that this could be the explanation for a scatter of this magnitude. Moreover, the scatter appears to be approximately as wide as the scatter in Table 10, for which the arguments that the scatter is due to respondent inaccuracies are more persuasive. This leads to the conclusion that while the amounts calculated from the pension documents are too high, much of the scatter in Table 11 is due to respondent misreporting.

Given this conclusion, it would appear that the amounts calculated from the pension

documents, while too high, are probably better guides to the accumulations than are the respondent reports. In this situation, the best approach may be to take the values calculated from the pension documents and apply a correction to reduce the apparent bias. Table 14 presents the results of a median regression attempting to quantify this overestimation for the sample of plans with voluntary contributions. The dependent variable is the ratio of the accumulations reported by respondents to the value calculated from the plan documents, which is the amount plotted on the vertical axis in Figure 1. The independent variable is the log of the pension value calculated from the plan documents, and its square. This value is the variable plotted on the horizontal axis of Figure 1. The estimated function has a value of approximately unity at a pension value of \$1,000 and drops to 0.48 for a pension value of \$25,000 and 0.36 for a pension value of \$100,000. To correct for the apparent bias in the values calculated from the pension documents, we will adjust those values by reducing them according to the results implied by the regression in Table 14.²⁶

Projecting Pensions from Linked Data for Those Without a Linked Employer Record

The last task is to predict the value of the pension, adjusted by the procedure described in the last paragraph, based only on the information in the respondent's survey. This will allow us to predict the values of pensions for respondents for whom the survey was not able to obtain pension documents. At first glance, it would seem that we could take the respondents' defined benefit and defined contribution pensions and project them separately, and then combine the results.

However, the problem is complicated by the fact that a substantial number of respondents are

²⁶ Regressions with additional variables were tried, but in these regressions only a small number of variables were significant, and they suffered from the problem that they yielded negative predicted values for some of the pensions.

misinformed about the type of plan they have, as shown in Table 3. For example, for a respondent who reported only a defined benefit plan, we could regress the values calculated from the plan documents on the values calculated from the respondents' expected benefit amounts, but this would miss the fact that of respondents who report only a defined benefit plan, almost half in fact either have an additional defined contribution plan or have only a defined contribution plan.

There are two ways to deal with this problem. One is to use multinomial logits or probits to predict the probability of a respondent actually having a defined benefit, defined contribution, or both types of plans conditional on his answer to the question about plan type on the survey. Then, conditional on each of these plan types, we could predict the value of the defined benefit and defined contribution (as appropriate) conditional on the pension values calculated from information that the individual reports. This would require, for instance, calculating the probability that an individual who reports a defined benefit plan in fact has a defined contribution plan, and then inferring the value of the defined contribution plan from the reported defined benefit plan plus additional variables. In the end, we would take the expected value of the imputed values of the various types of plans weighted by the probability that the individual has them, conditional on the respondents' reported plan types and pension values.

A simpler way, and the one that will be pursued here, is to combine all the pensions that the individuals have and to try to impute the combined value from the information available from the respondents. To do so, however, we must be careful about combining the pension amounts. The value of the defined benefit plan is the value that the respondent expects to collect and includes the effects of future years of work. On the other hand, the value of the defined contribution plan is the amount that is accumulated in the respondent' account, which presumably

only includes the effects of work to the date of the survey. To combine the two amounts for a single individual, we should either extend the defined contribution amount out to the respondent's retirement date or only consider the part of the defined benefit value that is due to work to date.

The choice is largely dictated by the purpose of the exercise. To compute a pension amount that reflects current wealth as of the survey date, adjusting the defined benefit amount would be appropriate. To compute a final pension wealth amount at retirement, adjusting the defined contribution amount forward to the expected retirement age would be more advisable. Here, we construct a pension amount to be used in conjunction with other components of wealth as of the survey date, so we adjust the defined benefit amounts to correspond to the amount as of the survey date. Unfortunately, there is no unique way to do this, since two approaches are available. One approach figures the value of the pension using current tenure and current earnings, and assumes that the individual separates from the firm immediately. A second approach figures the discounted value of the pension assuming that the participant works until the expected retirement age, and prorates the pension based on the proportion of eventual years of service already accumulated. In the literature, the former approach is known as the "legal" value of the pension, while the latter is known as the "projected" value of the pension. The argument is akin to the familiar problem of allocating the price of an asset over time: there are a variety of methods of figuring depreciation. In the current context, there is probably an argument for using the projected measure of pension value, since the legal measure effectively ignores the option value that the participant has acquired by working at the firm to date.

To calculate the respondent's total pension value, we rely on two amounts, one based on the pension documents, and one based only on the respondents' responses. The procedures are

similar to each other except for defined contribution pensions with voluntary contributions. The accumulations directly reported by the respondents are taken as is, but the amounts calculated from the pension documents are reduced according to the procedure discussed above to allow for increasing contributions over time. To the accumulations for DC plans with voluntary contributions, modified as described, we add the value of defined contribution pensions without voluntary contributions. This gives two figures for total defined contribution pensions: one based on plan documents, and one on accumulations reported by respondents. For defined benefit plans, we multiply the amount calculated from the expected benefit amounts by the fraction of years until retirement that have already been served. Then the total defined benefit amount is added to the total defined contribution amount, separately for the amounts calculated from plan documents and for the amounts based on the respondents' answers. Thus, it would be possible for the total pension value, calculated from the plan documents, to be the sum of a defined benefit pension and a defined contribution pension with voluntary contributions, if those pensions are indicated in the plan documents. At the same time, the total pension value based on the respondent's answers would simply be a defined contribution pension with no voluntary contributions, if the respondent listed only that type of pension when asked about pensions.

Table 15 tabulates the resulting total pension values as calculated from plan documents vs. the total values calculated only from respondents' information.²⁷ Note that unlike the previous tables, this one includes cases where the respondent told the survey that he or she had one type of pension but the plan documents indicated another type. The table looks reasonably symmetric

²⁷Note that the value of DC plans with voluntary contributions is adjusted according to the coefficients in Table 14.

around the main diagonal, as would be expected since we have eliminated systematic discrepancies between calculated values and reported accumulations for the defined contribution plans. However, there is a wide scatter, indicating that there remains a large difference between the amounts calculated from the plan documents and those reported by the respondents.

Table 16 reports on regressions to explain the values calculated from the plan descriptions, which are taken to be approximately correct.²⁸ In total there are six regressions. The first four regressions use observations in which pension value estimates are available from both the provider and the respondent. The first and third of these regressions use only the value of the plan calculated from the survey questions as an explanatory variable. The second and fourth add additional explanatory variables. The second and fourth regressions use the preferred specification, and these results give some idea as to how much adding in the additional explanatory variables improves the fit of the regression. The fifth and sixth regressions explain the value of pensions obtained from provider plan descriptions, but pertain to observations where pension amounts are not available in the respondent survey. Since the final two regressions use only the observations for which the pension value cannot be calculated from the survey questions, it performs does not include that value. From Table 10 we saw that respondents who did not report amounts for their pensions had, on average, lower values calculated from the plan documents than did respondents who did report a value. Hence, it would not be advisable to take

²⁸Note that the provider plan values for DC plans with voluntary contributions in Table 16 are adjusted according to the coefficients in Table 14. The regressions include variables for missing values for plan characteristics, with the value for the missing variable indicator set to 1 if the characteristic is relevant to the type of plan and the plan characteristic is missing, and the value of the missing variable set to zero, so that the value is picked up in the coefficient of the variable indicating the value is missing.

a mechanical approach, applying a regression for those who did report pension values to impute pensions for those who did not report pension values without any further adjustment.

Equations 1 and 3 and equations 2 and 4 are related regressions. Because the distributions of pension values appear to be roughly loglinear, if regressions were run on the linear values computed from the plan documents vs. the linear values calculated from the survey questions, the regressions would give enormous weight to very high value pensions. To avoid this, we take the log of the pension values. However, this creates another problem, since some of the pensions, especially those calculated from the plan documents, have zero values. For the explanatory variable (the value calculated from the survey questions), we can take care of this by creating a binary variable which takes on a value of one if the pension value is zero. For the dependent variable, it requires estimating the pension value in two steps. The first step is a probit for the probability that the pension value calculated from the plan documents will be zero. As might be expected, this probability declines rapidly as the pension value calculated from the survey questions increases. The second step is a regression of the log of the value from the plan documents on a set of explanatory variables, conditional on the value being positive. The probits are in equations 1, 2 and 5; the regressions are in the third, fourth and sixth equations.

For those who have positive employer-provided and respondent pensions, equation 3 suggests an elasticity of employer pension value with respect to reported respondent value of around .73. The R2 for that equation is 0.63, suggesting that self reported plan value is associated with about two thirds of the variation in the plan value calculated from firm reports. Holding other plan features and job characteristics constant, equation 4 suggests an elasticity of firm reported plan values with respect to respondent values of .24, but that holds constant a

number of plan characteristics that are associated with higher plan value. Thus equation 4 should be used to predict pension values, but a great deal of care should be exercised in attempting to interpret any particular coefficients.

Other Sources of Error:

We have not yet considered another source of error on the part of respondents. Some individuals may not correctly report whether they are covered by a pension on their job. One worrisome scenario occurs when a respondent reports the firm does not offer a pension, but the firm does in fact offer a pension and the respondent is covered. When informed by the respondent that the firm does not offer a pension, the HRS does not attempt to collect a plan description. One way to search for such an error is to compare responses in wave 1 and wave 2 for respondents who said they were in the same job. If the respondent said he had a pension in wave 1, he was asked about changes in that pension in wave 2, wherein he could deny he had a pension. If the respondent said he did not have a pension in wave 1, he is asked in wave 2 whether he has a pension. As seen from Table 17, about 20% of those saying they did not have a pension in 1992 reversed their stance in 1994, but relatively few who said they did have a pension in 1992 denied it in 1994. Whether the asymmetry is due to the differences in the questions asked in 1994, or whether the respondent reports a pension only if he decides to participate, and those reporting coverage decided to participate in '94, is open to question. Given that a denial is stronger than simply responding afresh to the pension question, we would have expected this pattern to some degree, but perhaps not to this extent.

C. Social Security Comparisons

The Social Security earnings records provided by the HRS are unique to this group of

surveys. Only a few researchers who work at or with the Social Security Administration have access to comparable data. From these records, we can closely estimate for HRS respondents the amount of covered earnings, computed as Average Indexed Monthly Earnings (AIME), and values of their Social Security benefits, computed as the Primary Insurance Amount (PIA). Although there is a small chance of a mismatch, the earnings records are a gold standard. Therefore patterns of discrepancies between respondent reports and the benefits computed from the earnings histories are very informative. Moreover, these relationships are useful in predicting the values of Social Security benefits for the 25 percent of HRS respondents for whom we do not have Social Security earnings records.

Table 18 summarizes the distributions of reported earnings based on Social Security records and respondent reports of earnings histories for the 9472 respondents for whom HRS has a matched earnings record.²⁹ The data in the table are AIME computed on an annual basis, i.e., AIME multiplied by 12. Using respondent reports, including earnings on the current or last job in 1992, at the start of the current or last job, in a previous job, in past pension covered jobs, and incorporating information from wave 3 on the age the respondent initially entered the labor force, years of full time work, and years of covered work, we overestimate average indexed yearly earnings by about 6.6 percent, overestimate average yearly earnings for men by about 3.5 percent, and overestimate average indexed yearly earnings for women by about 23 percent. When the

²⁹Table 18 includes the social security records obtained based on permissions granted by respondents in wave 1, as analyzed in Table 1, and in addition it includes social security records obtained with permissions granted by respondents in waves 2 and 3.

differences are ordered from low to high, however, the median difference is greatly reduced.³⁰

Table 19 displays the joint distribution of the Average Indexed Monthly Earnings on a yearly basis. The dispersion is narrower than the dispersion of pension values. In Table 19, there are a total of 9472 observations; of those, 7303, or 77 percent, fall within one cell of the diagonal, and 4066 observations, or 43 percent of the observations fall exactly on the diagonal. The R^2 is 0.82.

The univariate distributions of the primary insurance amounts on an annual basis, as computed from respondent earnings reports and from SSA earnings records are reported in Table 20. The means as computed from respondent earnings reports are 5.8 percent higher than the means computed from Social Security earnings records. For men, the values computed from respondent reports are 1.4 percent above the amounts computed from Social Security earnings records, and for women, using the values computed from respondent reports of earnings causes benefits to be overstated by 13.5 percent. It is likely that the cap on covered earnings plays some role in generating the discrepancy, but the cap cannot account for the fact that benefits for women are overstated by so much more than the overstatement found for men.

³⁰It would also be of interest to estimate the relation of the benefit calculated from the respondent's earnings record and the benefit the respondent expects, as reported in question N46 of the HRS. However, looking at the correlation between the respondent's estimated social security amount and the amount calculated from the records runs against several problems. First, less than half of the respondents ventured a guess as to the amount. Looking at the counts for question N46, 5815 respondents reported that they expected to receive social security benefits, but only 2563 ventured an amount. Secondly, since the question was in section N, the financial respondent answered the question for both spouses, meaning that there is a non-random sample of respondents answering the question for themselves. Third, there is a temporal mismatch: the amount constructed from the record pertains to earnings through 1991, but the amount self-reported presumably includes the effects of any future work. In addition, it is reported in dollars of some future year, and we don't have a good fix on the price levels in that year.

In Table 21, which involves more cells than in Table 19, 6033 observations, or 64 percent of the observations, fall within one cell of the diagonal. The observations along the diagonal account for 31 percent of all observations (2942/9472).

Tables 22 and 23 report the relations between the present values of own benefits as computed from respondent earnings reports and those obtained from the Social Security Administration. Overall, the present value of own benefits is seen in Table 22 to be 5.9 percent too high when computed from respondent reported earnings. For men the lifetime benefits are about 0.8 percent too high when computed from respondent earnings histories, and for women they are 13.6 percent too high. From Table 23, we find that 68 percent of the observations are within one cell of the main diagonal (6428/9472), and that the observations along the main diagonal account for 34 percent of the observations (3251/9472).

Table 24 reports coefficients for equations that relate the AIME on an annual basis, computed from the Social Security earnings record, to variables that are observable in the survey. These equations include observables associated with the availability of an earnings record, and may be used to predict the annualized AIME for those observations without a Social Security record on the basis of the observable characteristics for those respondents. When separate equations are estimated for men and women, the regression accounts for 60 percent of the variance in annualized AIME for men, and 66 percent of the variance for women. Since this equation is meant for prediction, rather than for analytical purposes, it includes a number of different measures of earnings and benefits. Consequently, the coefficients on particular variables are not readily interpretable.

IV. Conclusions

We began the study by asking whether it is possible to do a reasonable job in explaining who has a linked Social Security or pension record and who does not. There is some systematic relationship. Those with the least valuable pensions are less likely to have a pension match. Those with higher incomes are less likely to have a pension or Social Security match. This creates a problem for researchers who would base their analysis of pension or Social Security wealth only on a sample that includes respondents with linked data. Nevertheless, observables do not do a very good job in explaining who gave permission to obtain a Social Security record. Pension matches are more predictable, but the explanatory power is still not very high. Thus in neither case is there is a very strong role for observables, suggesting that the penalty from restricting the sample may not be very high. However, we have not generated evidence on selectivity with respect to unobservables.

We have also investigated how pension plan features and values differ under two circumstances: when the information is obtained directly from respondents, and when the information is obtained from the employer. In addition, we compared covered earnings histories developed from respondent reports, and the implied Social Security benefits, with covered earnings histories obtained from the Social Security Administration, and the implied benefits.

For those who have matched records, we also ask whether it is possible to explain the variation in pension values and earnings histories based on employer-provided pension records or Social Security administrative data, using information from respondent reports, including respondent perceptions of pension values or earnings histories, to provide the explanations. Relatively good explanations are obtained. For most purposes, these results suggest that it will be

possible to predict pension and Social Security values for those without a matched record based on the relationship between respondent reported outcomes and administrative or firm-provided data for those with a record available. The equations we estimate can be used to predict pension and Social Security values for those in the HRS without an attached pension or Social Security record. The accompanying files we produce for HRS users can be used not just to impute pension and social security wealth levels for those without a matched record, but can also be used by researchers who require information on pension or social security wealth levels, but do not wish to or cannot obtain permission from the HRS to use restricted data.

Nevertheless, observed discrepancies between the firm provided and administrative records on the one hand, and respondent reports on the other are large for many respondents, large enough to suggest that a great deal of caution is required in any behavioral research that uses pension or Social Security measures as explanatory variables. In addition, in the case of the dates of early and normal retirement and the benefit reduction rates associated with defined benefit plans, the discrepancies between self reported values and values obtained from linked data are too wide to permit imputation. Thus researchers will have to be especially cognizant of the effects of reporting error on parameter estimates, especially when attempting to predict the effects of nonlinear benefit formulas on the retirement decisions made by those who do not have a pension plan description available.

The implications of this situation depend on the nature of the analysis. For the estimation of retirement models, particularly models whose purpose is to estimate the effect of economic incentives on retirement, researchers may be justified in omitting observations if the Social Security record, or more importantly the pension record, is not available. Given the magnitudes

of the errors in reporting early and normal retirement ages in defined benefit plans and benefit reduction rates, it seems likely that the retirement incentives will be severely misrepresented if only the respondent reports can be used. Many respondents may retire at ages that correspond to their true incentives, but if those incentives are inferred by the researcher to occur at other ages, it will appear to the researcher that respondents are not responding to economic incentives. The situation is better in studies that seek to simulate behavior under different hypothetical scenarios. In simulations, it is not quite so crucial that pensions be matched exactly with the right individuals, as long as the statistical matches are valid.

Our findings also suggest that the behavioral models may have to be modified. It appears that older respondents today are pessimistic about what their plans will be worth. This pessimism is recent, and different from the situation ten years ago, when there was an important subgroup of respondents who was overly optimistic about its benefits, and when the respondents would be permitted to retire. The wide discrepancies we find at the individual level between the pension reports of respondents and firms, and the Social Security amounts based on earnings reports from respondents and computed from earnings records, further suggest that even those on the verge of retirement have imperfect information about their pensions and Social Security. Many respondents approaching retirement may not be reacting to the correct information, but may be making their decisions based on a faulty understanding of their retirement programs. Accordingly, the discrepancies we find raise important questions about models of retirement and saving that assume that households are fully informed about their financial condition, and that they fully understand the incentives presented by complex pension and Social Security rules.

Table 1: Probit For Probability Permission Was Given in Wave 1 to Match the Social Security Earnings Record

Independent Variable	Partial Effect	z
female	0.000	0.01
single	-0.034	-2.94
no child	-0.018	-0.99
child not known	-0.173	-4.81
black	-0.028	-2.26
Hispanic	-0.061	-3.96
high school drop out	0.014	1.22
some college	-0.021	-1.69
college graduate	-0.028	-1.70
some graduate school	-0.041	-2.50
home owner	-0.008	-0.56
assets \$0 to 10K	0.075	4.02
assets 10K to 25 K	0.072	3.64
assets 25 K to 100K	0.035	2.99
assets 250K to 1000K	-0.038	-3.00
assets 1000+K	-0.010	-0.44
assets not known	-0.190	-4.12
retirement horizon <2 years	-0.001	-0.08
retirement horizon 2 to 4 years	0.006	0.33
retirement horizon 10+ years	0.001	0.09
retirement horizon never	-0.040	-2.29
retirement horizon not applicable	-0.042	-2.21
retirement horizon not asked from proxy	-0.540	-23.64
retirement horizon not known	-0.047	-2.70
firm size 0 to 4 employees	0.021	1.04
firm size 5 to 14 employees	0.008	0.38
firm size 15 to 24 employees	0.012	0.38
firm size 25 to 99 employees	0.039	1.88
firm size 500+ employees	0.029	2.01
firm size not known	-0.013	-0.54
nonmanufacturing	-0.002	-0.17
industry not known	-0.035	-0.62
management	-0.010	-0.89
white collar	-0.015	-1.19
occupation not known	-0.172	-1.84
union	0.020	1.49
union not known	0.054	2.38
annual earnings \$0 to 15K	-0.015	-0.74
annual earnings \$15 to 30K	-0.029	-1.51
annual earnings \$30 to 50K	-0.036	-1.90
annual earnings \$100+K	-0.035	-0.96
annual earnings not known	-0.069	-2.02
pension	-0.009	-0.76
pension status not known	0.184	1.95
cons		8.26
Number of Observations, 12,652	Log Likelihood, -7158.599	Pseudo R2, 0.0652

Table 2: Probit for Probability There Is a Matched Pension Provider Survey

Independent Variable	All Pensions		Current Job Pension		Last Job Pension		Previous Job Pension	
	Partial Effects	z	Partial Effects	z	Partial Effects	z	Partial Effects	z
female	0.010	0.747	0.010	0.54	0.073	2.077	-0.011	-0.485
single	-0.023	-1.396	-0.018	-0.866	-0.044	-1.032	-0.036	-1.299
no child	0.014	0.542	0.004	0.107	0.015	0.237	0.051	1.287
child not known	-0.032	-0.663	-0.100	-1.619	-0.071	-0.614	0.093	1.180
black	0.053	3.099	0.000	0.014	0.107	2.677	0.103	3.541
Hispanic	-0.020	-0.775	-0.041	-1.261	-0.123	-1.887	0.053	1.146
high school drop out	-0.076	-4.505	-0.034	-1.528	-0.114	-3.237	-0.102	-3.429
some college	0.033	2.061	0.007	0.328	0.070	1.667	0.046	1.748
college graduate	0.082	3.887	0.087	3.043	0.133	2.278	0.042	1.301
some graduate school	0.097	4.722	0.120	4.373	0.174	2.762	0.042	1.280
home owner	0.050	2.349	0.018	0.627	0.133	2.463	0.055	1.612
assets \$0 to 10K	0.030	1.131	0.018	0.499	0.113	1.675	0.017	0.393
assets 10K to 25 K	-0.009	-0.324	-0.021	-0.569	-0.022	-0.306	0.030	0.645
assets 25 K to 100K	-0.001	-0.079	0.017	0.871	-0.031	-0.811	-0.004	-0.171
assets 250K to 1000K	-0.005	-0.28	0.026	1.162	-0.049	-1.218	-0.010	-0.358
assets 1000+K	-0.114	-3.178	-0.106	-2.08	-0.116	-1.335	-0.113	-2.069
assets not known	-0.051	-0.65	0.060	0.603	0.157	0.663	-0.338	-1.943
retirement horizon <2 years	0.062	3.447	0.010	0.298	0.136	1.623	0.136	4.659
retirement horizon 2 to 4 years	0.036	1.683	0.043	1.77	0.075	0.521	0.024	0.601
retirement horizon 10+ years	-0.019	-1.117	-0.013	-0.65	0.142	1.423	-0.026	-0.864
retirement horizon never	-0.037	-1.383	-0.040	-1.203	0.049	0.366	-0.011	-0.270
retirement horizon not applicable	0.005	0.122			0.130	1.369	-0.009	-0.124
retirement horizon not asked from proxy	-0.012	-0.386	-0.042	-1.14	0.129	1.186	0.042	0.821
retirement horizon not known	-0.013	-0.492	0.005	0.147	-0.119	-0.953	0.010	0.239
tenure 0 to 1	0.013	0.367	-0.027	-0.685	0.029	0.345	-0.197	-1.202
tenure 2 to 4	0.024	0.989	-0.039	-1.36	0.083	1.271	0.089	1.498
tenure 10+	0.058	3.806	0.016	0.736	0.058	1.227	0.077	3.402
tenure not known	-0.138	-1.156	-0.333	-1.506	-0.124	-0.503	0.029	0.140
firm size 5 to 14 employees	-0.409	-9.053	-0.391	-8.325	-0.421	-4.083	n.a.	
firm size 15 to 24 employees	-0.254	-4.736	-0.283	-4.863	-0.165	-1.509	n.a.	
firm size 25 to 99 employees	-0.238	-8.855	-0.248	-8.633	-0.193	-3.425	n.a.	
firm size 500+ employees	0.014	0.876	0.012	0.675	0.025	0.752	n.a.	

Independent Variable	All Pensions		Current Job Pension		Last Job Pension		Previous Job Pension	
firm size not known	-0.221	-3.984	-0.227	-3.735	-0.074	-0.547	n.a.	
nonmanufacturing	0.091	5.571	0.088	3.98	0.055	1.405	0.069	2.594
industry not known	0.032	0.458	0.043	0.452	-0.090	-0.583	0.071	0.607
management	-0.110	-7.091	-0.047	-2.272	-0.141	-3.671	-0.141	-5.468
white collar	-0.013	-0.75	-0.031	-1.381	-0.075	-1.832	0.040	1.455
occupation not known	-0.253	-2.076	-0.062	-0.329			-0.362	-2.093
union	0.006	0.379	0.044	2.482	-0.007	-0.216	n.a.	
union not known	-0.120	-2.134	-0.169	-0.949	0.235	0.845	n.a.	
annual earnings \$0 to 15K	0.022	0.772	0.062	1.631	-0.109	-1.467	0.051	1.012
annual earnings \$15 to 30K	-0.007	-0.282	0.025	0.786	-0.105	-1.52	0.007	0.131
annual earnings \$30 to 50K	-0.001	-0.027	0.028	0.946	-0.125	-1.869	0.051	0.993
annual earnings \$100+K	-0.307	-5.348	-0.285	-4.593	-0.347	-1.557	-0.281	-1.970
annual earnings not known	-0.079	-1.042	-0.047	-0.447	0.224	1.008	-0.179	-1.354
dc plan	-0.185	-5.96	-0.190	-4.807	-0.074	-0.91	-0.165	-2.570
combination plan	-0.044	-1.669	-0.055	-1.884	0.131	1.615	-0.072	-1.160
Plan type not known	-0.226	-5.159	-0.189	-3.147	-0.070	-0.572	0.016	0.069
DB plan: annual benefits 0 to 2K	-0.089	-3.44	-0.073	-1.591	-0.151	-2.6	-0.037	-1.013
DB plan: annual benefits 2 to 5K	-0.065	-2.345	-0.087	-2.084	-0.074	-1.348	-0.009	-0.200
DB plan: annual benefits 10 to 25K	0.104	4.346	0.026	0.778	0.080	1.648	0.209	5.293
DB plan: annual benefits 25+K	0.143	4.536	0.152	3.409	-0.027	-0.423	0.203	3.839
DB plan: annual benefits dk	0.005	0.19	-0.016	-0.5	-0.078	-1.067	-0.191	-0.851
DC plan: 0-10K in account	0.047	1.704	0.085	2.686	-0.093	-1.101	0.032	0.518
DC plan: 10-25K in account	0.070	2.288	0.073	2.178	0.141	1.386	-0.012	-0.163
DC plan: 100-250K in account	0.088	1.884	0.084	1.704	-0.033	-0.226	0.150	1.093
DC plan: 250+K in account	0.147	2.153	0.132	1.717	0.199	0.769	0.134	0.927
DC plan: dk amount in account	0.055	1.964	0.064	2.034	-0.079	-0.977	0.056	0.860
constant		1.578		1.716		0.159		-4.399
No. Of Obs.	8682		4448		1384		2839	
Pseudo R2	0.1574		0.1164		0.1355		0.1221	
Log Likelihood	-5031.335		-2539.922		-772.28188		-1606.1744	

Table 3: Pension Plan Type as Reported by the Respondent and the Firm, for Current Job Held in 1992, Including Only Those Respondents with a Matched Pension Plan

3A: Self Reported vs. Firm Provided Plan Types

Self Reported					
Provider Report	DB	DC	Both	DK	Total
DB	777	213	380	22	1392
DC	160	327	111	10	608
Both	409	163	315	20	907
Total	1346	703	806	52	2907

3B: Percentages with Self Reported vs. Firm Provided Plan Types

Self Reported					
Provider Report	DB	DC	Both	DK	Total
DB	27.0%	7.0%	13.0%	1.0%	48.0%
DC	6.0%	11.0%	4.0%	0.0%	21.0%
Both	14.0%	6.0%	11.0%	1.0%	31.0%
Total	46.0%	24.0%	28.0%	2.0%	100.0%

3C: Percentages with Self Reported Plan Type Conditional on Firm Report of Plan Type (Percent of Row Total)

Self Reported					
Provider Report	DB	DC	Both	DK	Total
DB	55.8%	15.3%	27.3%	1.6%	100.0%
DC	26.3%	53.8%	18.3%	1.6%	100.0%
Both	45.1%	18.0%	34.7%	2.2%	100.0%
Total	46.3%	24.2%	27.7%	1.8%	100.0%

3D: Percentage with Firm Reported Plan Type Conditional on Respondent Report of Plan Type (Percent of Column Total)

Self Reported					
Provider Report	DB	DC	Both	DK	Total
DB	57.7%	3.0%	47.1%	42.3%	47.9%
DC	11.9%	46.5%	13.8%	19.2%	20.9%
Both	30.4%	23.2%	39.1%	38.5%	31.2%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

Table 4: Self Reported vs. Firm Provider Reported Dates of Early Retirement for Plans Reported as Defined Benefit by Both the Respondent and the Firm

Provider Reported	Self Reported																			Total
	<50	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	>65	DK	
<50	19	5	5	3	5	2	54	1	2	2		7		23		1	2		7	138
50	7	24	6	3	4	1	65	3	4	4	5	8	1	14		1	1		17	168
51	3	6		2	1		5	1	2	2	2	3		5			1		3	36
52	1	3	2				11		1		1	7		1					4	31
53	2		1	8	1	1	9				2	3	1	1			1		2	32
54				3	1	2	8	1		1	2	5		4			1		3	31
55	15	11	4	6	9	12	314	22	22	22	28	37	2	160	5	4	46	7	43	769
56	3	2		2			20	2	2	1	1	6	2	9	1		5	2	4	62
57	3		1	1			11	4		3	3	6	2	9	2		1		7	53
58		1				1	6	1	8	6	3	3	1	9			3		4	46
59	1				2	1	11		3	4	3	9	1	6			5		6	52
60	9	5	1	3	1	6	26	8	4	7	8	40	5	68		2	17	1	25	236
61		1					4	1		1		7	4	7	1		1		2	29
62	2	1			1		13	1	3	3	3	7	3	28	2	2	11	1	7	88
63							1					1		6	2			2	1	13
64	1						1		1			1		8	2		1		3	18
65							2							21	4	1	8		8	44
>65							1	2				1		3	1		5	1	3	17
DK								1		2		1		8			2	1	3	18
Total	66	59	20	31	25	26	562	48	52	58	61	152	22	390	20	11	111	15	152	1881

Table 5: Self Reported vs. Firm Provider Reported Dates of Early Retirement for Plans Reported as Defined Benefit by Both the Respondent and the Firm, Including Only Observations Of Wave 1 Respondents Who Plan to Retire By 1995

Provider Reported	Self Reported																			DK	Tot
	<5	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	>6			
<50	6	2	1	2	1		17	1				3		5						2	40
50	1	5			2	1	17		2	1	1	2		1		1				2	36
51	1	3					2					1		2						1	10
52		1					4		1			1								2	10
53			1				3				1	1								1	8
54						1	4				1	3									8
55	4	3	1	2	3	1	57	2	5	1	5	10	1	27	2	1	4	1	5	135	
56	1	1					9					2			1			1		1	15
57	2						2	1					2	1	1					1	10
58							2	1	3	1		1		1						1	10
59					1		1					1								1	4
60	1	1		2		3	8		1		2	10		10		1	3			3	45
61							2					3		1						1	7
62							2		1		1		1	6				1		1	13
63												1		2	1						4
64							1					1		1							3
65							1							4	1			1		3	10
>65												1		1							2
DK																			1		1
Total	16	16	4	6	7	6	132	5	13	3	11	41	4	62	6	3	9	3	24	371	

Table 6: Self Reported vs. Firm Provider Reported Dates of Normal Retirement for Plans Reported as Defined Benefit by Both the Respondent and the Firm

Provider Reported	Self Reported																			Total
	<50	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	>65	DK	
<50	1	1	2		2		6				1	3		5			1			22
50		7					6	1	1			5		7	2		7		5	41
51		1		1			1					1		1			2		1	8
52	1	1	1	1			2				1	2		2		1			1	13
53				4	1	1	5		1	2		1							1	16
54	1			1	1	1	3	1				1		2			1		1	13
55	4	2	1	1	2	2	81	2	3	1	3	5		15		2	13	2	7	146
56				1			9	3	1	1	1	1		3			3	2	1	26
57							13	4	7	2		2		2	1		4	1	1	37
58		1			1	1	8	2	5	4	1	1		5		1	6		2	38
59	1				1		11		3	4	3	4		3	1	1	4		3	39
60	1	2	1	1		4	63	6	9	5	15	69	6	67	7	2	50	9	26	343
61							4	1	1			13	2	5			5	2	2	35
62	5	2	1	3	2		19	4	5	7	8	13	10	95	10	4	68	16	21	293
63							2		1		1	1		5	1	1	3	1	3	19
64					1		3			2	1	1		1	2	1	2	1	2	17
65	5	2	2	4	3	3	74	3	12	15	21	41	7	188	6	7	287	17	42	739
>65												1		3	1		4	6	2	17
DK	1									1		1		6		1	6	1	2	19
Total	20	19	8	17	14	12	310	27	49	44	56	166	25	415	31	21	466	58	123	1881

Table 7: Self-Reported vs. Provider Annual Early Retirement Reduction Rates, for Plans Reported as Defined Benefit by Both the Respondent and the Firm

	Self Reported											Total
Provider Reported	0-1%	1-2%	2-3%	3-4%	4-5%	5-6%	6-7%	7-8%	8-10%	>10	% NA	Total
0-1%	1	9	4	6	1	4	4			5	102	136
1-2%	5	3	3	2	2	4	4	4	1	6	83	117
2-3%	1	3	2	5		8	2	6		4	97	128
3-4%	2	2	4	7		10	3	2	1	3	123	157
4-5%	3	7	4	5	3	5	4	4	2	9	191	237
5-6%	4	1	3	1	1	6	4	2		7	102	131
6-7%				1		1			1		47	50
7-8%		1		1					1		27	30
8-10%	2			2		2	1		1	2	31	41
>10%		1		1		2	3			3	62	72
DK/NA	7	28	14	20	2	16	7	9	6	26	647	782
Total	25	55	34	51	9	58	32	27	13	65	1512	1881

Table 8: Self-Reported vs. Provider Voluntary Contributions

	Self Reported		
Provider Report	No	Yes	Total
No	121	57	178
Yes	303	392	695
Total	424	449	873

Table 9: Distribution of Pension Values for Current Jobs Held at Time of the Survey (\$1992)

	Defined Benefit Plans		Defined Contribution Plans	
	Respondent Reported	Provider Reported	Respondent Reported	Provider Reported
Percentile				
95	385,497	523,704	200,000	347,265
90	311,775	387,276	128,322	223,951
75	211,733	227,373	45,000	100,030
50	116,327	112,380	15,000	29,067
25	48,084	49,767	4,500	6,093
10	16,805	24,223	1,200	750
5	9,231	11,252	400	0
Mean Values	148,015	168,405	59,105	85,790
Number of Observations	1,122		641	

Columns 1 and 2 are present values as of the date of expected retirement, discounted to 1992. They are not prorated to allow comparisons between individual and firm reports of DB plan values, as the individual reports were as of the date of expected retirement. The DB and DC amounts in this table are not comparable since the DC amounts are account balances as of 1992, and do not include future contributions.

Table 10: Self Reported Vs. Provider Reported Amounts Accumulated in Defined Benefit Plans

Provider Reported	Self Reported											Total
	0	0-5K	5-10K	10-20K	20-50K	50-100K	100-200K	200-500K	500- 1M	>1M	DK	
0		6	3	3	5	4	6	5	1		157	190
0K-5K			2	1	1			1			7	12
5K-10K		3	2	2	1	1	3	2			6	20
10K-20K		5	13	7	6	3	3	1			42	80
20K-50K		3	19	31	65	38	25	9	1		133	324
50K-100K		2	3	8	51	91	55	14			141	365
100K-200K	1	1	5	3	21	63	131	62	1		113	401
200K-500K				6	7	16	84	147	2		63	325
500K-1M		1			3		5	43	7	2	6	67
> 1M								2	2			4
DK		1			5	2	2	1			82	93
Total	1	22	47	61	165	218	314	287	14	2	750	1881

Table 11: Self Reported Vs. Provider Reported Amounts Accumulated in Defined Contribution Plans

Provider Reported	Self Reported											Total
	0	0-5K	5-10K	10-20K	20-50K	50-100K	100-200K	200-500K	500- 1M	>1M	DK	
0	2	28	8	9	6	4	2				38	97
0K-5K	1	58	14	7	5	1	1				25	112
5K-10K		22	13	5	2	1		1			12	56
10K-20K	1	18	22	23	13	3	2				26	108
20K-50K	1	18	19	21	37	10	2	2	1		31	142
50K-100K		9	12	13	27	18	15	3			35	132
100K-200K		5	5	11	27	16	12	6	2	1	21	106
200K-500K		1	1	3	14	15	14	14	2	3	12	79
500K-1M						1		4			1	6
> 1M			1			1		1	1			4
DK		4	1	4	11	3					51	74
Total	5	163	96	96	142	73	48	31	6	4	252	916

Table 12: Self Reported Vs. Provider Reported Amounts Accumulated in Defined Contribution Accumulated Amounts for Plans without Voluntary Contributions

Provider Reported	Self Reported											Total
	0	0-5K	5-10K	10-20K	20-50K	50-100K	100-200K	200-500K	500-1M	>1M	DK	
0	1	10	3	1	1	1					11	28
0K-5K		15	4	2	2	1	1				9	34
5K-10K		4	4	3	1	1					2	15
10K-20K		4	2	6	5	2	2				6	27
20K-50K		3	4		8	2	2	2	1		8	30
50K-100K		2		2	2	3	6	2			10	27
100K-200K		2	1		3	1	1	2	1	1	4	16
200K-500K					1	1		1		2		5
500K-1M												
> 1M												
DK				1	1						4	6
Total	1	40	18	15	24	12	12	7	2	3	54	188

Table 13: Reported Accumulations and Calculated Values of Defined Contribution Pensions

	Calculated Value of Pension								
	0-5K	5-10K	10-20K	20-50K	50-100K	100-200K	200-500K	500K-1M	>1M
	Pensions Without Voluntary Contributions								
Median Accumulation Reported By Respondents	1,900	7,000	14,000	24,850	92,000	40,000	400,000	-	-
Median Value Calculated from Pension Documents	1,207	7,668	15,924	27,972	67,050	142,295	287,970	-	-
Number of Observations	42	13	21	22	17	12	5		
	Pensions With Voluntary Contributions								
Median Accumulation Reported By Respondents	3,375	4,200	7,000	15,000	30,000	37,500	94,500	206,000	201,750
Median Value Calculated from Pension Documents	1,146	6,945	15,090	30,389	69,141	131,885	292,690	550,933	1,151,112
Number of Observations	104	31	61	89	80	73	62	5	4

Table 14: Median Regression for the Over Prediction of Pension Values Calculated from Plan Documents

Defined Contribution Plans with Voluntary Contributions

Dependent variable is the ratio of the value of the pension calculated from plan documents to the accumulation reported by the respondent.		
Explanatory Variables:	Coefficient	t-statistic
Constant	3.1822	3.56
ln (pension value calculated from plan documents)	-0.4200	2.39
[ln (pension value calculated from plan documents)] ²	0.0152	1.77
Number of Observations	467	

Table 15: Pension Value As Calculated From Plan Documents Vs. Pension Value As Calculated From Respondent Reports, All Plans

Self Reported												
Provider Reported	0	0-5K	5-10K	10-20K	20-50K	50-100K	100-200K	200-500K	500-1M	>1M	DK	Total
0	5	67	30	35	31	14	11	4	2	0	295	494
0K-5K	1	77	20	7	11	1	2	1	0	0	88	208
5K-10K	0	26	24	22	14	2	5	0	0	0	81	174
10K-20K	1	20	26	43	49	18	4	2	0	0	98	261
20K-50K	0	11	12	38	92	61	31	14	2	0	176	437
50K-100K	0	3	4	9	39	90	72	21	2	0	163	403
100K-200K	0	1	5	2	15	59	121	56	4	2	111	376
200K-500K	0	1	0	4	4	15	67	128	9	1	53	282
500K-1M	0	0	0	0	2	0	3	26	6	2	6	45
> 1M	0	0	0	0	0	0	0	2	2	1	0	5
DK	1	4	2	2	10	4	3	1	0	0	143	170
Total	8	210	123	162	267	264	319	255	27	6	1214	2855

Table 16: Regression of Provider Total Pensions on Self-Reported Pensions Plus Other Variables

Independent Variable	Dependent Variable =1 if Provider Pension = 0		Dependent Variable =1 if Provider Pension = 0		Dependent Variable is In Provider Pension Value; Includes Observations With Positive Provider and Self Reported Values		Dependent Variable is In Provider Pension Value; Includes Observations With Positive Provider and Self Reported Values		Dependent Variable = 1 if Provider Pension = 0; Includes Observations With Zero Self Reported Values		Dependent Variable is In Provider Pension Value; Includes Observations With Positive Provider and Zero Self Reported Values	
	Coefficient	z	Coefficient	z	Coefficient	t	Coefficient	t	Coefficient	z	Coefficient	t
constant	1.8518	-3.634	1.3436	0.843	2.8820	17.516	0.3946	.627	1.1667	0.609	-1.7049	-1.791
respondent reports zero pension value	-1.5584	-12.302	-0.1049	-0.199	6.2749	14.304	2.1122	5.941				
ln of value respondent reports	-0.3005	7.610	-0.1265	-3.044	0.7347	48.866	0.2393	11.914				
current age			-0.0001	-0.012			0.0641	13.126	-0.0103	-0.847	0.0925	14.631
age at hire			0.0051	0.839			-0.0588	-24.723	0.0268	4.348	-0.0780	-23.671
age expects benefits to begin			0.0277	0.020			-0.0304	-4.114	-0.0482	-2.486	-0.0124	-1.514
age benefit expected missing									-0.6440	-0.535	-1.6050	-3.096
respondent has only dc plan			2.1980	1.712			-2.7741	-6.090	0.0819	0.062	-3.0848	-5.504
respondent has db and dc, or combination plan			-0.2497	-1.107			-0.2605	-3.290	-0.4248	-1.829	-0.2903	-2.542
early retirement age, db plan, =0 for DC plan			0.1299	0.666			0.0036	0.599	0.0153	0.836	-0.0272	-3.468
respondent reports can collect benefits at any age							0.0180	0.036			-1.7266	1.921
early retirement age missing			0.6858	0.570			0.1531	0.422	1.0917	0.989	-1.7534	-3.694
normal retirement age, db plan, =0 for DC plan			-0.0191	-1.134			-0.0037	-0.695	0.0087	0.486	0.0016	0.205
normal retirement age missing			-1.5753	-1.395			-0.1793	-0.529	0.5958	0.529	0.1609	0.324
annual reduction factor, db plans, =0 for dc plans			-3.5662	-0.954			-0.0411	-0.080	1.4351	0.901	-0.0061	-0.008
reduction factor missing			-0.2735	-0.966			-0.0267	-0.411	0.0914	0.353	0.1601	1.407
contribution rate for 401k/403b/sra plans, 0 otherwise			-1.0265	-0.704			1.2385	1.981	-3.8001	-2.039	1.7229	1.964
contribution rate missing			-0.1614	-1.008			0.1486	1.941	0.2449	1.208	0.3791	3.326
female			0.0170	0.142			0.1380	3.035	-0.3595	-2.658	0.2552	3.703
ln earnings			-0.2999	-2.819			0.7975	16.156	-0.1231	-0.923	1.2024	16.621
zeroearn			-2.8484	-2.144			8.2435	13.437			(dropped)	
fulltime			-0.1308	-0.834			-0.1273	-1.924	-0.2316	-1.319	-0.3411	-3.531
union			0.2526	2.234			0.0259	0.632	0.1450	1.268	0.1835	3.112
manufacturing			0.2150	1.326			-0.0450	-0.685	-0.2963	-1.392	-0.2656	-3.018

Independent Variable	Dependent Variable =1 if Provider Pension = 0		Dependent Variable =1 if Provider Pension = 0		Dependent Variable is In Provider Pension Value; Includes Observations With Positive Provider and Self Reported Values		Dependent Variable is In Provider Pension Value; Includes Observations With Positive Provider and Self Reported Values		Dependent Variable =1 if Provider Pension = 0; Includes Observations With Zero Self Reported Values		Dependent Variable is In Provider Pension Value; Includes Observations With Positive Provider and Zero Self Reported Values	
	Coefficient	z	Coefficient	z	Coefficient	t	Coefficient	t	Coefficient	z	Coefficient	t
public employment			0.3303	2.019			0.4032	6.931	-0.0031	-0.018	0.4075	4.130
manager or professional			0.3655	2.359			0.2653	4.632	0.4016	2.168	0.0191	0.216
white collar			-0.0198	-0.135			0.2453	4.394	0.5067	3.043	0.0576	0.707
firm size >100			0.0586	0.386			-0.0494	-0.738	-0.4179	-2.472	0.0589	0.529
firm size missing			0.7342	1.794			0.0237	0.114	-0.1610	-0.348	-0.1085	-0.296
health is good			-0.0766	-0.701			0.0066	0.156	0.1376	1.129	0.0144	0.227
health is poor			0.0467	0.255			-0.0348	-0.474	0.0621	0.328	-0.0761	-0.788
plans over next few months or years			0.1812	1.545			-0.0019	-.04	0.0711	0.541	-0.0063	-0.090
planning horizon is 5 to 10 years, or more than 10 years			0.0591	0.348			0.0687	1.013	-0.0121	-0.057	-0.0145	-0.144
planning horizon not available			0.2955	0.915			0.2286	1.621	-0.3081	-1.074	-0.0578	-0.377
number of words recalled in second test			-0.0163	-0.873			0.0167	2.427	-0.0119	-0.559	-0.0008	-0.072
missing recall measure			-0.0247	-0.079			-0.0393	-0.289	0.1249	0.393	0.1580	0.952
Number of Observations	1614		1611		1415		1415		1067		776	
Adjusted or Pseudo R2	0.1525		0.2686		0.6297		.7970		.4133		.7256	

Table 17: Pensions Over Time In The Same Job

		1994 Observation	
		No Pension	Pension
1992 Observation	No Pension	1075	271
	Pension	73	3138

Table 18: SSA Vs. Respondent Reported Earnings (AIME, Expressed on an Annual Basis, in 1992 Dollars)

	All Respondents			Male Respondents			Female Respondents		
	Respondent	SSA	Difference	Respondent	SSA	Difference	Respondent	SSA	Difference
Mean	13378	12548	-1190	19909	19230	-680	8545	6926	-1620
Percentile									
95	31958	30820	7720	33341	32587	10221	24526	20228	4702
90	29422	28252	4785	32022	31034	6915	20330	16211	2950
75	22922	20773	1481	28356	27261	2609	13280	10425	778
50	12110	10070	-453	21866	20551	-245	6554	5067	-611
25	3969	3513	-3660	12366	11570	-3778	1850	1754	-3572
10	462	903	-8225	3853	4772	-8815	0	394	-7730
5	0	264	-11952	757	2180	-13526	0	91	-10945
Number of Obser.	9472			4328			5144		
Correlation	0.82			0.72			0.76		

This table includes the Social Security records obtained with permissions granted by respondents to HRS waves 1, 2 and 3. Average Indexed Monthly Earnings, expressed on an annual basis, equals AIME*12.

Table 19: Distribution of Respondent And SSA Reports of Average Indexed Yearly Earnings (Thousands of 1992 Dollars)

Respondent Report	SSA Records											Total
	0 - 4	4 - 8	8 - 12	12 - 16	16 - 20	20 - 24	24 - 28	28 - 32	32 - 36	36 - 40	40+	
0 - 4	1803	358	105	34	27	21	18	14	2	0	0	2382
4 - 8	395	519	182	70	32	16	10	4	2	0	0	1230
8 - 12	173	338	358	136	49	25	10	6	2	0	0	1097
12 - 16	72	148	264	228	144	63	22	9	0	0	0	950
16 - 20	47	62	145	217	178	125	75	18	3	1	0	871
20 - 24	31	40	63	117	195	184	136	41	10	0	0	817
24 - 28	32	13	46	50	94	206	287	125	16	0	0	869
28 - 32	13	20	14	23	39	76	188	334	83	0	0	790
32 - 36	6	6	16	14	24	25	44	141	172	0	0	448
36 - 40	0	1	0	0	0	1	3	3	4	3	0	15
40+	0	0	0	0	1	0	1	0	1	0	0	3
Total	2572	1505	1193	889	783	742	794	695	295	4	0	9472

This table includes the Social Security records obtained with permissions granted by respondents to HRS waves 1, 2 and 3. Average Indexed Yearly Earnings equals Average Indexed Monthly Earnings, expressed on an annual basis, which equals AIME*12.

Table 20: SSA Vs. Respondent Reported Primary Insurance Amount, Expressed on an Annual Basis (in 1992 Dollars)

	All Respondents			Male Respondents			Female Respondents		
	Respondent	SSA	Difference	Respondent	SSA	Difference	Respondent	SSA	Difference
Mean	6293	5948	-345	8399	8282	-117	4521	3984	-537
Percentile									
95	12003	11806	3153	12189	12087	3829	10078	8780	2659
90	11507	11156	1944	12014	11844	2404	8823	7607	1561
75	9574	8940	623	11201	10867	887	6725	5861	464
50	6371	5764	-171	9252	8873	-59	4727	4209	-295
25	3323	2959	-1316	6431	6203	-1124	1554	1479	-1495
10	386	758	-2901	3201	4012	-2716	0	332	-3053
5	0	222	-4216	646	1833	-4208	0	76	-4219
Number of Obser.	9472			4238			5144		
Correlation	0.81			0.71			0.76		

This table includes the Social Security records obtained with permissions granted by respondents to HRS waves 1, 2 and 3. The Primary Insurance Amount is multiplied by 12.

Table 21: Distribution of Primary Insurance Amount Based on Respondent And SSA Reports (Thousands of 1992 Dollars, Expressed on an Annual Basis)

Respondent Report	SSA Records														Total
	0 - 1	1 - 2	2 - 3	3 - 4	4 - 5	5 - 6	6 - 7	7 - 8	8 - 9	9 - 10	10-11	11-12	12-13	13+	
0 - 1	731	240	99	64	67	45	16	14	9	10	9	10	1	0	1315
1 - 2	120	128	84	43	54	25	12	5	5	3	1	1	0	0	481
2 - 3	60	88	93	62	62	33	12	5	5	2	5	4	0	0	431
3 - 4	42	49	65	71	87	36	8	6	2	4	2	1	0	0	373
4 - 5	54	70	112	121	250	133	50	31	19	9	6	3	1	0	859
5 - 6	46	45	52	70	221	291	125	49	29	11	6	4	1	0	950
6 - 7	17	24	30	32	137	209	195	103	63	26	14	6	2	0	858
7 - 8	16	10	20	19	46	114	180	145	104	54	29	11	2	1	751
8 - 9	12	7	11	8	31	72	112	137	131	95	63	18	6	0	703
9 - 10	11	5	10	9	21	41	53	101	150	135	109	50	8	0	703
10-11	14	3	4	9	3	27	33	48	87	149	220	120	17	0	734
11-12	4	6	1	6	16	14	17	26	50	88	159	375	76	0	838
12-13	3	1	3	0	7	9	15	10	20	23	44	150	176	0	461
13+	0	0	0	1	0	0	0	2	1	1	3	2	4	1	15
Total	1130	676	584	515	1002	1049	828	682	675	610	670	755	294	2	9472

This table includes the Social Security records obtained with permissions granted by respondents to HRS waves 1, 2 and 3. The Primary Insurance Amount is multiplied by 12 to express it on an annual basis.

Table 22: SSA Vs. Respondent Present Value of Social Security Benefits (1992 Dollars)

	All Respondents			Male Respondents			Female Respondents		
	Respondent	SSA	Difference	Respondent	SSA	Difference	Respondent	SSA	Difference
Mean	55745	56222	-3123	71880	71305	-844	42170	37130	-5041
Percentile									
95	113163	109625	28315	119057	117684	32868	97495	85268	24281
90	100862	97570	17501	110794	108923	20649	83782	72367	14584
75	81624	76572	5558	94583	92596	7374	62816	54192	4128
50	55678	50471	-1423	75379	72692	-400	42207	36548	-2585
25	29682	26637	-11635	51592	49702	-9327	13963	13511	-13789
10	3331	6807	-26340	26403	30988	-23195	0	2977	-27956
5	0	2031	-38845	4891	14643	-36082	0	665	-40012
Number of Obser.	9472			4328			5144		
Correlation	0.81			0.75			0.77		

This table includes the Social Security records obtained with permissions granted by respondents to HRS waves 1, 2 and 3.

Table 23: Distribution of Present Values of OWN Social Security Benefits Based on Respondent And SSA Reports
(1,000's of 1992 Dollars)

Respondent Report	SSA Records														Total
	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100	100- 110	110- 120	120- 130	130 +	
0 - 10	788	249	104	73	69	30	17	13	9	6	6	6	1	1	1372
10 - 20	137	151	92	71	40	20	8	5	2	3	0	1	0	0	530
20 - 30	78	101	126	91	52	20	6	4	3	2	2	2	1	0	488
30 - 40	54	91	112	225	130	53	18	13	8	3	2	2	0	1	712
40 - 50	57	63	95	213	343	132	70	20	21	5	0	3	1	0	1023
50 - 60	32	44	46	98	285	286	157	58	23	12	2	4	1	0	1048
60 - 70	24	19	18	52	131	210	227	138	50	20	3	5	0	1	898
70 - 80	20	12	22	30	73	118	195	231	125	40	16	1	3	2	888
80 - 90	10	1	9	13	30	60	102	204	206	100	29	9	3	1	777
90 - 100	7	8	7	13	25	35	53	927	157	260	63	17	2	4	743
100-110	3	5	1	6	9	7	12	23	50	103	155	38	4	1	417
110-120	9	0	1	4	5	7	8	15	19	20	67	127	23	8	313
120-130	1	2	1	0	1	4	1	6	4	7	13	29	52	7	128
130+	3	0	1	1	0	4	3	4	1	4	10	11	19	74	135
Total	1223	746	635	890	1193	986	877	826	678	585	368	255	110	100	9472

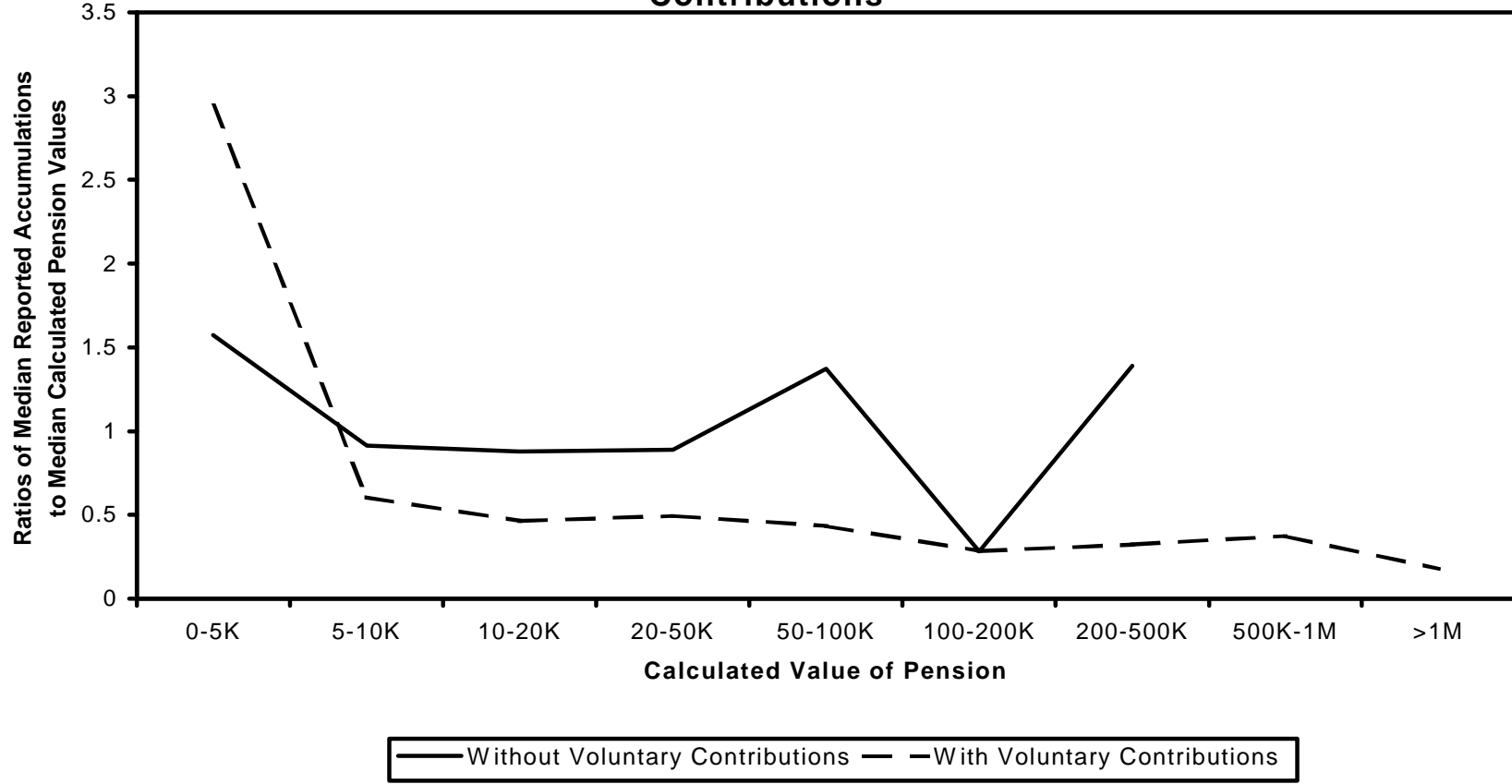
This table includes the Social Security records obtained with permissions granted by respondents to HRS waves 1, 2 and 3.

Table 24: Regressions of Average Indexed Yearly Earnings Calculated from Social Security Records on Average Indexed Yearly Earnings Calculated from Earnings Imputed From Survey

Independent Variables	All Observations With Matched Social Security Earnings Records				Observations for Males With Matched Social Security Earnings Records		Observations for Females With Matched Social Security Earnings Records	
	Coefficient	t	Coefficient	t	Coefficient	t	Coefficient	t
constant	1862.567	-19.016	4981.976	13.855	3402.957	3.648	1398.607	4.305
AIME calculated from earnings imputed from survey	0.7777729	137.799	0.508322	45.367	0.496085	26.184	0.464773	32.268
AIME imputed from expected benefit amount			0.098246	12.103	0.098774	7.976	0.081594	7.976
AIME from expected benefit missing			575.1495	2.844	-140.401	-0.296	551.66	2.785
AIME from earnings imputed from survey interacted with missing value indicator			-0.01288	-1.123	0.022119	1.061	-0.02161	-1.405
female			-5007.04	-37.722	(dropped)		(dropped)	
married			133.7257	0.983	1533.768	5.438	-702.969	-5.434
respondent working at time of survey			460.8062	1.822	919.6561	1.078	543.8686	2.509
earnings in current job			0.037378	7.207	0.021982	2.789	0.050138	7.390
indicator respondent had section G job			-618.953	-2.284	-1119.76	-1.27	-305.116	-1.291
earnings from section G job			0.04359	6.015	0.047313	4.047	0.065221	6.634
self employed			-1206.42	-7.305	-1151.8	-4.161	-1337.15	-7.311
worked for federal government in past			-1304.03	-6.934	-2234.92	-7.254	-148.641	-6.98
worked for a state government in past			-262.562	-2.006	-312.554	-1.388	-141.215	-1.040
reported a nonSocial Security job in past			-991.027	-6.688	-1186.12	-4.735	-1025.29	-6.330
current or last five year job was fulltime			-399.307	-2.710	-744.787	-2.319	-102.916	-0.742
respondent had a pension in any job			2441.358	17.927	3051.091	12.179	2068.543	14.877
union in main job			150.4643	1.088	303.0848	1.331	-37.2376	-0.237
management/professional in main job			1099.166	8.029	1308.136	5.657	990.0998	6.561
manufacturing in main job			1640.825	11.643	1860.837	8.255	1183.972	7.164
indicator survey available for 1996			35.57308	0.215	698.9399	2.506	-801.524	-4.402
reported health excellent or very good			105.3806	0.849	152.5965	0.692	142.7883	1.110
reported health fair or poor			-595.914	-3.873	-725.832	-2.643	-456.815	-2.866
health variable not available			(dropped)		(dropped)		(dropped)	
plans over next few months or next year			-332.302	-2.74	-577.749	-2.582	-96.458	-0.792
plans over 5 to 10 years, or > 10 years			-155.735	-0.813	135.3838	0.399	-374.559	-1.888
planning horizon not available			-448.932	-1.5	-404.633	-0.772	-605.107	-1.887
words recalled in second test			33.78321	1.735	47.17687	1.254	39.51627	2.082
recall variable missing			415.102	1.282	509.4233	0.937	433.8685	1.162
Number of Observations	9472		9472		4328		5144	
Adjusted R2	0.6672		0.7535		0.5992		0.6550	

This table includes the Social Security records obtained with permissions granted by respondents to HRS waves 1, 2 and 3. Average Indexed Yearly Earnings equals AIME times 12.

Ratios of Median Reported Accumulations to Median Calculated Pension Values for DC Plans With and Without Voluntary Contributions



References

- Anderson, Patricia, Alan Gustman and Thomas Steinmeier. 1999. "Trends in Labor Force Participation and Retirement: Some Evidence on the Role of Pensions and Social Security in the 1970's and 1980's". *Journal of Labor Economics* 17(4), pt. 1, October: 757-783.
- Bernheim, B. Douglas. 1988. "Social Security Benefits: An Empirical Study of Expectations and Realizations". In R.R. Campbell and E. Lazear, editors, *Issues in Contemporary Retirement*. Stanford, California: Hoover Institution. 312-335.
- Coile, Courtney and Jonathan Gruber. 1999. "Social Security and Retirement". Paper presented at NBER Summer Institute.
- Gustman, Alan L. and F. Thomas Juster. 1996. "Income and Wealth of Older American Households: Modeling Issues for Public Policy Analysis". In Eric A. Hanushek and Nancy Maritato, editors, *Assessing Knowledge of Retirement Behavior*. Washington, D.C.: National Academy Press: 11- 53.
- Gustman, Alan L., Olivia S. Mitchell, Andrew A. Samwick and Thomas L. Steinmeier. 1999. "Pension and Social Security Wealth in the Health and Retirement Study". In James Smith and Robert Willis, editors, *Wealth, Work and Health, Innovations in Measurement in the Social Sciences*. Ann Arbor: University of Michigan Press: 150-208.
- _____, forthcoming. "Evaluating Pension Entitlements". In Olivia Mitchell, editor, *Forecasting Retirement Needs and Retirement Wealth*. Philadelphia: University of Pennsylvania Press for the Pension Research Council.
- Gustman, Alan L. and Thomas L. Steinmeier, 1989. "An Analysis of Pension Benefit Formulas, Pension Wealth and Incentives from Pensions", *Research in Labor Economics* 10: 53-106.
- _____, 1999a. "Employer-provided Pension Data in the NLS Mature Women's Survey and in the Health and Retirement Study". National Bureau of Economic Research, Working Paper 7174. June.
- _____, 1999b. "Changing Pensions in Cross-Section and Panel Data: Analysis with Employer-provided Plan Descriptions"
- Haider, Steven and Gary Solon. 1999. "Evidence of Non-Response Bias in the HRS Social Security Files". Xerox. Ann Arbor, Michigan.
- Lumsdaine, Robin. 1996. "Factors Affecting Labor Supply Decisions and Retirement Income". In Eric A. Hanushek and Nancy Maritato, editors, *Assessing Knowledge of Retirement Behavior*. Washington, D.C.: National Academy Press: 61-115.

Lumsdaine, Robin L. and Olivia S. Mitchell. Forthcoming. “New Developments in the Economic Analysis of Retirement”. In Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*.

Olson, Jan. Forthcoming. “Linkages with Data from SSA Records in the Health and Retirement Study”. *Social Security Bulletin*.

Appendix Table 1: Matching Rates of Social Security Records and Pension Plan Descriptions
With Respondent Records

	Matching Rates for Social Security Records	Number of Total Observations	Matching Rates for Pension Plan Descriptions	Number of Observations
Gender				
Men	71.1	[5867]	55.5	[4972]
Women	72.2	[6785]	54.6	[3710]
Combined	71.6	[12652]	55.2	[8682]
Age				
<= 50	73.7	[1595]	60.7	[971]
51-61	71.3	[9742]	54.7	[6951]
>= 62	72.4	[1315]	57.7	[1008]
Marital Status				
Married	72.4	[9896]	56	[6997]
Single	70.1	[2756]	53.4	[1685]
Parents				
Yes	72.2	[11743]	55	[8003]
No	68.3	[754]	57.9	[547]
Race				
White	71.8	[9415]	54.7	[6870]
Black	71.7	[2064]	62.7	[1360]
Hispanic	69	[1173]	48.4	[452]
Education				
< HS	73.4	[3696]	47.6	[1675]
HS Grad	74.1	[4424]	54.5	[3064]
Some Coll	69.4	[2320]	55.7	[1812]
Coll Grad	68.6	[1040]	56.3	[951]
Grad Sch	66.7	[1172]	62.4	[1180]
Residence Status				
Homeowner	71.1	[10205]	56.2	[7517]
Renter	73.3	[2447]	49.6	[1165]
Assets				
< 10K	75.8	[1949]	49.9	[781]
10-25K	77.5	[812]	51.8	[447]
25-100K	74.6	[3565]	56	[2571]
100-250K	70.8	[3304]	57.7	[2744]
250K-1M	65.2	[2393]	57.5	[1809]
1M+	63.9	[534]	33.1	[280]
Retirement Horizon				
< 2 Yrs	73.5	[2939]	62.7	[2146]
2 to 4 years	74.2	[1040]	61.6	[972]
4 to 9 years	74.5	[2071]	55.6	[1794]

	Matching Rates for Social Security Records	Number of Total Observations	Matching Rates for Pension Plan Descriptions	Number of Observations
10+	74.1	[2396]	50.1	[2077]
Never Retire	71.1	[1069]	45.5	[549]
Self Employed				
Yes	67.9	[1564]	0	[184]
No	72.8	[10022]	56.8	[8498]
Combined	72.1	[11586]	55.2	[8682]
Tenure with Firm				
< 2 Years	72.6	[578]	51.7	[269]
2 to 4 Years	71.2	[775]	52.5	[657]
5 to 9 Years	73.8	[2286]	41.7	[1825]
10+ Years	71.7	[7944]	60	[5894]
Size of Firm				
< 5	71	[805]	17	[90]
5-14	72.4	[609]	25.3	[163]
15-24	65.4	[242]	45.4	[97]
25-99	74.6	[611]	40.3	[432]
100-499	75.2	[750]	56.5	[710]
500+	72.7	[1634]	73.9	[1784]
Industry				
Manufacturing	74.7	[2520]	49.5	[2318]
Other	71.5	[8971]	57.5	[6283]
Occupation				
Mgmt/Prof	69.2	[3318]	57	[3006]
White Collar	70.7	[2871]	52.6	[2113]
Blue Collar	75.4	[5356]	55.2	[3528]
Union Status				
Union	72.1	[2122]	70.1	[2253]
Nonunion	71.9	[5448]	62.1	[3576]
Annual Earnings				
< 15K	74.4	[4630]	44.2	[2427]
15-30K	71.9	[3519]	54.6	[3169]
30-50K	70.3	[2089]	63.5	[2069]
50-100K	72.4	[879]	67.2	[810]
100K+	63.5	[191]	30.4	[128]
Pension				
Yes	72.6	[5056]		
No	71.8	[6530]		

The observations in this table use the HRS survey weights. Percentages in column 2 are for permissions as reported in V136, wave 1. Job characteristics in column 1 are for the longest job. Calculations for columns 3 and 4 include only jobs offering pensions. Numbers in brackets are the observations over which the percentage is taken. Percentages are weighted averages.

Appendix Table 2: Matched Pension Plans by Type of Pension and Pension Value

Respondent Reported Pension Characteristic	Matching Rates for Pension Plan Descriptions	Number of Observations
Pension Type		
Defined Benefit	59.1	[4528]
Defined Contrib	40.8	[2446]
Combination	69.2	[1368]
DB Annual Benefit		
< 2K	41.8	[664]
2-5K	49.8	[663]
5-10K	61.4	[794]
10-25K	71.6	[1475]
25K+	76.9	[642]
DC Account Value		
< 10K	46.9	[1258]
10-25K	58.9	[627]
25-100K	54.5	[641]
100-250K	68.1	[184]
250K+	52.0	[86]

The observations in this table use the HRS survey weights.