

PRELIMINARY, COMMENTS WELCOME

**What You Don't Know Can't Help You:
Knowledge and Retirement Decision Making**

September 2003

Sewin Chan
Wagner Graduate School of Public Service
New York University
sewin.chan@nyu.edu

Ann Huff Stevens
Department of Economics
University of California, Davis
annstevens@ucdavis.edu

Abstract

This paper focuses on a puzzle raised by two strands of the recent retirement literature: how can individuals respond to detailed pension information that they, apparently, do not possess? Recent evidence shows that most individuals do not know the details of their employer-provided pension plans. At the same time, virtually all recent studies of retirement timing are based on administrative data of which individuals themselves may not be aware. We show that individuals do respond to pension incentives constructed from self-reported data even after conditioning on incentives constructed from employer-reported data. We use the self-reports to construct measures of individual knowledge about pension plans and show that individual responsiveness to pension incentives based on employer-reported data vary dramatically with these knowledge measures. Well-informed individuals have an elasticity of retirement with respect to pension incentives that is three times as large as the average response when information issues are ignored. In contrast, the uninformed show no evidence of a relationship between their pensions and retirement timing. Finally, we show that our findings seem robust to a selection model, and that there is no evidence that endogeneity of information acquisition fully explains the heterogeneity in responsiveness to financial incentives. These findings should encourage researchers to make more use of self-reported data to better understand decision-making.

The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) to the Center for Retirement Research at Boston College (CRR). This grant was awarded through the CRR's Steven H. Sandell Grant Program for Junior Scholars in Retirement Research. The opinions and conclusions are solely those of the authors and should not be construed as representing the opinions or policy of SSA or any agency of the Federal Government or of the CRR. Helpful comments were received from seminar participants at Syracuse University.

I. Introduction

Recent empirical research has found that the typical individual approaching retirement is not well-informed about the details of his or her private pension plan. At the same time, a large body of research suggests that individuals are quite responsive to the financial incentives embedded in these pension plans. Virtually all of this analysis of responsiveness to financial incentives is based on administrative data that individuals may or may not be aware of. This paper focuses on the puzzle raised by these two strands of literature: how can individuals respond to fairly detailed information that they do not, apparently, possess? We argue that, while typical estimates of the effects of pensions based on administrative or employer-reported data may accurately capture an average effect, this average masks a tremendous amount of heterogeneity in individual responsiveness to financial retirement incentives that is directly related to their knowledge of pension details. We utilize self- and administratively-reported pension and Social Security data from the Health and Retirement Study to explore the importance of individual knowledge in determining the response to pension-related financial incentives for retirement.

A better understanding of differential responsiveness to pension incentives that is related to individuals' knowledge of these incentives is crucial if the estimated effects of pensions on retirement are to be used to inform current or future policy discussions. Several recent studies in this area (Samwick (1998), and Coile and Gruber (2000), for example) use their estimates to simulate or forecast the response to various policy changes, such as altering the normal retirement age or benefit levels. Because alternative policies (and the specific implementation of policy changes) will almost certainly be associated with different levels of transparency and public knowledge, understanding how this knowledge interacts with the financial incentives

themselves is crucial to accurately predicting what the response will be, or at a minimum, which types of individuals are most likely to respond to a given policy.

A possible explanation for the puzzle posed above is that earlier studies showing ignorance of pension details simply reflect severe measurement error in self-reported pension data. Perhaps individuals do know and understand their pensions, but self-reported data on pensions are so inaccurate that they provide a misleading picture of individual knowledge. We present evidence that this is unlikely to be the case. While it is almost certainly true that self-reported pension data are measured with some error, financial incentive measures based on these self-reported data are significant predictors of retirement behavior, even after conditioning on incentive measures derived from administrative data.

A key contribution of this paper is to demonstrate that self-reported pension data, while largely ignored by economists studying retirement behavior, do provide relevant information. This is an important finding given the extensive effort devoted to collection of self-reported pension data, and the rarity with which those data have been used by economists to study retirement behavior. It is of practical importance to establish whether and how this extensive data can be used in studies of retirement behavior. If the self-reported data truly are so poorly measured as to be uninformative, that would raise the question of whether they should continue to be collected. However, as we illustrate below, the self-reported pension data are useful in understanding retirement decision-making.

This paper uses self-reported, employer-reported and Social Security Administration data from the Health and Retirement Study (HRS) to examine the relationship between knowledge of retirement benefits and individuals' responsiveness to those benefits. We begin by calculating two versions of financial retirement incentives, one based on self-reported data, and another

based on employer-reported or administrative data. It is straightforward to show that both versions of these incentives affect retirement, and that the self-reports continue to be strongly significant determinants of retirement even after controlling for the employer reports. We then use the self-reports to construct summary measures of individuals' knowledge of their pension plans. When these knowledge measures are interacted with the incentives based on employer-provided data, we show that the better-informed fraction of the population generates all of the average response to pension incentives, and that the ill-informed segment of the population show no significant response to those incentives.

As noted above, our investigation of how people respond to the self- and administratively- reported retirement incentives draws upon two strands of literature. The first of these includes studies that examine the extent and accuracy of information that individuals possess about their future retirement benefits. The common finding across these studies is that such information is typically quite incomplete. Most directly related to this paper, Gustman and Steinmeier (2001a, 2001c) examine determinants of knowledge concerning both Social Security and private pensions using the HRS, and the implications of this knowledge for retirement planning. They find, for example, that only half of respondents correctly identify the type of pension plan (defined benefit versus defined contribution) in which they participate, and that certain retirement planning activities, union membership and longer planning horizons signal individuals who are better informed about their pensions. In addition, knowledge seems to improve with age, or as workers approach typical retirement ages. Gustman and Steinmeier (2001c) also examine whether misinformed workers retire earlier or later than they had expected and find that workers who initially overestimate their available benefits retire later than they had

originally planned. However, they do not explore how the responsiveness to financial retirement incentives varies based on the extent of workers' knowledge.¹

The second strand of literature relevant to the current study includes virtually all of the existing literature examining the effects of private pension incentives on retirement and much of the related literature on Social Security incentives. Many papers have utilized data taken from employer-reports or administrative sources to examine the effects of pensions or Social Security on retirement decisions. Stock and Wise (1990), and Lumsdaine, Stock and Wise (1992) estimate both structural and reduced-form models of the retirement decision using data provided by a single large employer. Samwick (1998) uses the employer-provided data on private pensions from the Survey of Consumer Finances to estimate a regression-based counterpart to existing structural models of retirement, focusing on the role of forward-looking financial retirement incentives, or "option value" in determining the probability of retirement. In a similar approach, both Coile and Gruber (2000) and Gustman and Steinmeier (2001b) use employer-provided pension data, along with administrative records of Social Security earnings histories to estimate the effect of both private pension and Social Security benefits on retirement. None of these papers consider the effect of individual knowledge of these financial incentives on retirement behavior.

In our own previous work, we have deviated from this tendency to rely on administrative pension reports, and focused on exploring the role of pensions in retirement behavior and

¹ Other studies include Mitchell (1988) and Luchak and Gunderson (2000). Mitchell uses private pension information from the 1983 Survey of Consumer Finances provided by employers and workers and finds that membership in a union, job tenure, industry-level profitability, and the type of pension plan are related to the extent of workers' knowledge, with the strength of this relationship varying substantially across alternative knowledge measures. In addition, minorities and males are less fully informed than white and female workers. Luchak and Gunderson examine pension information among employees of a unionized public utility industry in Canada. They find that individuals in worse health, those with higher earnings, those who have already worked beyond their pension maximizing age, and those in professional or semiskilled professions are more likely to know key components of their pension plans.

retirement expectations using the self-reported pension data in the HRS (Chan and Stevens, 2002, 2003). We find the self-reported financial incentives to be important determinants of retirement behavior, even after conditioning on earlier retirement expectations, or after controlling for individual-specific fixed effects. The current paper confirms that both self- and employer-reported pension data are useful in understanding the determinants of individual retirement behavior.

II. Data and Samples

We use data from the first four waves of the Health and Retirement Study (HRS), including: (i) the self-reported, publicly released data on retirement income sources, retirement behavior, and retirement expectations; (ii) the employer-provided pension plan details, and (iii) the Social Security earnings histories, which allow us to calculate individuals' Social Security benefits at alternative future retirement dates. Several points should be made regarding these data sources.

First, the employer-provided pension data are available only for jobs held at or before wave 1 of the survey, while self-reported benefit data are available for all waves. Similarly, the administrative earnings histories include Social Security covered earnings only up to the initial wave. However, it is possible to use the employer-provided data to judge the accuracy of self-reported private pension information because job changes are not very frequent at this age range. Individuals who do change employers before retirement are dropped from our samples, as are the small fraction of individuals who report changes to their pension plans between waves.² The

² Individuals are asked whether there have been any changes to their pension plans since the previous survey. They are then asked about the details of their pensions, even if they report no change since the previous wave.

wage histories from Social Security records can be supplemented with self-reported earnings in subsequent years so that Social Security benefits can be calculated for later waves as well.

Because of this projection of the administrative data to subsequent waves, it will be important to control carefully for the survey wave in our analysis and compare results across the initial and later waves.

Second, we follow many previous researchers in assuming that the employer and Social Security administrative data represent the true benefit information. In some cases, particularly with respect to the employer reports, this assumption may not be accurate. For defined contribution pension accounts in particular, there are good reasons to expect that the employer reports could deviate from the truth (Uccello and Perese, 1999). Thus, in the case of defined benefit plans, we are more confident about relying on the employer reports as the “true” pension. Throughout the analysis, we examine the sub-sample of workers with defined benefit plans to investigate whether our results differ systematically across pension plan types.

Third, while the self-reported data are useful because they reflect what individuals actually know, many researchers have expressed concerns that such data may contain substantial measurement error. These pension reports may differ from the employer reports not only because individuals actually do not know the information, but also because they misreport this information. While this is certainly a concern, we believe that the potential importance of understanding how knowledge affects retirement decision-making overrides the concern about measurement error. Indeed, our preliminary results described below demonstrate that there are large and statistically significant differences between informed and uninformed individuals. These differences will be understated if our measures of knowledge are severely affected by measurement error.

Using the self- and employer-reported data on pensions, our first step is to take the pension plan details and use them to construct the present value of pension wealth available to the individual at each possible future retirement age. Following previous literature, we construct financial incentive measures that are based on the difference between: (i) the present value of pension wealth if an individual retires at the future date that maximizes his or her present value of pension wealth, and (ii) the present value of pension wealth if an individual retires today. We refer to this measure as an individual's "pension gain":

$$[1] \quad GAIN_{it} = PV(\text{pension wealth if } R = \text{argmax}_{(t+1 - \infty)} [PV(\text{pension wealth})]) \\ - PV(\text{pension wealth if } R=t)$$

where $GAIN_{it}$ is pension gain and R is the retirement date.³

For individuals self-reporting defined benefit pensions, we have information on annual benefit amounts, the normal retirement (pension eligibility) age, the early retirement age, cost of living adjustments, and benefit reduction amounts (in the case of early retirement) from each survey wave. For those reporting defined contribution pensions, we have information on the current account balance, employer- and own-contributions, and basic portfolio allocation decisions for existing DC accounts. These pieces of data are combined to generate estimates of the present value of pension wealth at each possible retirement age. The employer-provided reports, while more detailed, contain similar information for each pension plan. In calculating the present discounted value of pension wealth at each retirement age, we assume a discount factor of 3%, and weight future values by age- and gender-specific survival probabilities taken from Social Security Administration actuarial tables.

³This is similar to the option value concept used in Lumsdaine Stock and Wise (1992), and Samwick (1998), equivalent to the "peak value" used by Coile and Gruber (2000) and related to the "premium value" used by Gustman and Steinmeier (2001b).

Table 1 shows summary statistics for our sample. To be included in the analyses, individuals must have an employer-reported pension plan, and cannot yet be retired at the time of the wave 1 survey. Individuals remain in the sample through wave 4, or until they transition to retirement. Roughly 7 percent of all observations reflect a retirement in the current year, while 34 percent of individuals in our sample have retired by the time of the wave 4 survey. Slightly more than half of the sample is male, and the average age is just over 57 years. The sample includes 5 percent hispanic respondents and 16 percent black respondents. More than three-quarters of the sample members are married.

The middle section of Table 1 provides means for key summary measures of pension wealth, based on both self- and employer-reports. The average level of pension wealth (assuming the individual retires in the current year) is approximately \$111,000 (in 1992 dollars) based on self-reports, but is more than \$147,000 based on the employer-provided pension reports. The average pension gain, calculated as the present value of pension wealth if retirement occurs at the wealth-maximizing age minus the present value if retirement occurs today is \$84,392 from the self-reports and \$31,462 from the employer-reports. The mean incentive for continued work is somewhat higher in the self-reports. If we take the median values of pension gain, however, the relative values are reversed: the median pension gain based on self reports is \$0, while the median pension gain from employer reports is \$15,494.

The final section of Table 1 presents means of some variables intended to describe the accuracy of individuals' self-reported pension values. These measures are defined and discussed in the following section.

III. Empirical Strategy and Results

A. Comparing Self- and Administratively-Reported Financial Retirement Incentives

We begin with a direct comparison of the estimated effects of our summary measure of the financial incentives from private pensions (pension gain), calculated from both the employer reports and the self-reports. The first column of Table 2 shows the effects of pension gain from the employer-reports on the probability of retirement, using a simple linear-probability model and all observations for which there is an employer-reported pension available. As expected, the pension gain variable is negative and significant, implying that a larger gain to delaying retirement until a future optimal date reduces the likelihood of retirement. The estimated magnitude (-0.01) suggests that a one standard-deviation in pension gain (\$53,000) would result in a reduction in the retirement probability of roughly half a percentage point, or 8 percent.⁴ We also include a variable giving the present value of pension wealth if retirement occurs immediately. As in some earlier research, we find a positive and significant effect of the level of pension wealth⁵, suggesting that those with higher levels of pension wealth are more likely to retire. The magnitude of this coefficient is close to that reported in Coile and Gruber.

In the next column, we repeat the exercise, but restrict the sample to individuals who have valid self-reports of their pensions, as defined above. The coefficient on pension gain rises slightly, but is not significantly different from that in the first column.

⁴ The coefficient reported in Coile and Gruber (2000) on incentives from pensions and Social Security implies an effect that is approximately twice as large. In future work we will attempt to reconcile these differing estimates.

⁵ We should note that some other previous work has not found a significant effect of the level of pension wealth on retirement probabilities. Samwick (1998), for example, finds that the level of pension wealth has no effect on retirement once a forward-looking incentive measure is included in the model. Overall, however, these basic results are consistent both with theoretical predictions and with previous empirical work.

In the third column of this table, we use pension gain and pension wealth variables based on the self-reports. The coefficient on pension gain is somewhat smaller in absolute value, is within one standard error of the coefficient on the employer-reported pension gain from column two, and is itself statistically significant. In the fourth column, we include both the self- and employer-based versions of pension gain and pension wealth levels. The correlation between the two measures of pension gain is 0.22. When both are included together, the self-reported pension gain variable remains statistically significant and is somewhat larger in magnitude than the employer-reported variable. This is not consistent with the self-reported data being very noisy measures of the true pension reports that do not affect individual decision-making. For pension levels, the employer-reported data remain statistically significant, while the self-reported version does not.

In the final column of the table, we use the employer-reports to instrument for the self-reported data. This results in coefficients on the self-reported data that are approximately double in magnitude from the previous columns. This is likely partially explained by measurement error in the self-reported pensions. However, it may also reflect individual knowledge. If individuals who know that their self-reports are highly accurate are more responsive to those self-reports, the IV estimator should be higher since it will tend to more heavily weight individuals whose self- and employer-reports are highly correlated. This exercise demonstrates that there is clearly a role for the self-reported pension data. Individuals do appear to respond to their own beliefs about the structures and details of their pensions, even after conditioning on the true pension, as proxied by the employer-reports.

B. Exploring Heterogeneity in Responsiveness to Financial Incentives

Our next step in the empirical analysis is to document differences in the responsiveness of individuals to financial incentives based on a variety of measures of their understanding of these incentives. We want to capture the fact that there may be heterogeneity in the extent to which individuals respond to financial incentives, and that this heterogeneity may be related to the degree of information individuals possess about these incentives. Following much recent empirical work, we estimate a retirement equation described by:

$$[2] \quad R_{it} = \alpha_0 + \alpha_1 GAIN_{it} + \alpha_2 INFORMED_{it} + \alpha_3 GAIN_{it} * INFORMED_{it} + X_{it} \beta + \varepsilon_{it}$$

where R_{it} is a binary variable indicating that the individual i retires in year t , and ε is an error term. $GAIN_{it}$ represents the pension gain to continued work, as defined above, based on employer-reported pension data. Samwick (1998) and Coile and Gruber (2000) show that forward-looking measures such as pension gain have greater explanatory power than alternative myopic summary measures of financial status or incentives.

X_{it} is a vector that contains several important control variables, including age, demographic characteristics, calendar years, and measures of non-pension wealth and lifetime earnings.⁶ With the exception of the terms associated with coefficients α_2 and α_3 , this specification is identical to that used in previous work. Conceptually, $INFORMED_{it}$ is a variable that will reflect the extent of individuals' knowledge regarding their pensions and social security benefits. In their earlier work, Gustman and Steinmeier (2001a) construct several measures of this variable. We use several similar measures in our analysis, and also focus on measures that

come closer to capturing individuals' knowledge of the actual gain to continuing work (or of the forward-looking financial incentives used in recent retirement models). These are described along with estimates of [2] below.

In Table 3, we combine these measures of knowledge about pension plans with a linear probability model for retirement in which the key financial incentives to retire are summarized through the pension gain measure.⁷ Other variables included in these specifications are race, education, sex, self-reported health status, marital status, dummy variables for age, level of non-pension assets, a measure of average lifetime earnings, and dummies for the survey wave. In results not reported, we have also included indicators for retiree health insurance and an insurance indicator interacted with poor health. Adding these controls does not change our basic results; they are excluded here because of a larger number of missing values for the retiree health insurance variables. Standard errors have been corrected to account for repeated observations from the same individual across survey waves. We present results for men and women pooled, although we have also obtained similar results when we estimate separate regressions for men and women.

For ease of comparison, the first column of Table 3 replicates the results from the first column of Table 2 and shows the negative and statistically significant effect of pension gain on the probability of retirement. In the next columns of Table 3, we examine whether there is important heterogeneity in the responsiveness of individual retirement decisions to pension gain, and in particular whether such heterogeneity is related to the degree of information individuals

⁶ See Coile and Gruber (2000) for a discussion of the importance of carefully controlling for measures of lifetime earnings. It is also crucial to control fully for age since there are spikes and non-linearities in the age profile of retirement probabilities. Knowledge is also highly correlated with age. For these reasons we include a series of dummy variables for each year of age (51 to 70) in our sample.

⁷ All of our results have been replicated using a probit model, rather than a linear probability model, to estimate effects of retirement.

possess about their pensions. This requires specifying a definition of *INFORMED* from equation [2]. First, following the basic strategy of Gustman and Steinmeier (2001a), we consider whether individuals report “don’t know” when asked about the value of their pension. Specifically, we construct a variable, referred to as *REPORT*, that is equal to one for individuals who report enough of the key components of their pension plans to allow us to construct the present value of benefits at various possible retirement ages. If individuals report not knowing some major component of their pension, *REPORT* is set equal to 0. As noted by Gustman and Steinmeier, a large fraction of individuals report not knowing the value of their pension. As shown at the end of Table 1, 60 percent of our sample have *REPORT* equal to 1.

In column 2 of Table 3, we show that knowing (or at least reporting some estimate of) these key values is associated with an increased probability of retirement. Reporting these values increases the annual probability of retirement by approximately 2 percentage points. This may reflect reverse causality: individuals who are actively considering retirement may be more likely to know enough to report some values for their pension plan. Also of interest is the interaction between *REPORT* and pension gain. This interaction is negative and significant and larger in magnitude than the coefficient on pension gain itself in column 1, while the coefficient on pension gain itself is no longer significant.

An alternative measure used in previous work is to compare the present value of an individual’s *level* of pension wealth based on the self-report with that based on the administrative report.⁸ To implement this concept, we create a variable *KNOW_LEVEL* that is equal to one if (i) the ratio of the self-reported to employer-reported pension wealth is between 0.5 and 2, or (ii) one of the reported pension wealth values is zero, and the other is less than \$5,000 in absolute

⁸ Gustman and Steinmeier (2001), for example, create a variable equal to 1 if the self-report is within 25% of the administrative report of pension wealth.

value. The second component of this definition is included to avoid misclassifying as uninformed those who report zero wealth rather than very small amounts of pension wealth. For those who do not report enough information on their pension to calculate the value of pension wealth ($REPORT=0$), we set $KNOW_LEVEL=0$. We do this also for similar variables discussed below. Table 1 shows that 50 percent of our sample report levels of pension wealth are relatively close to those based on the employer reports. Column 3 of Table 3 displays results of this specification. The results again show that individuals with detailed knowledge of their pension plans show a greater responsiveness to pension-related incentives in making the retirement decision: the interaction of $KNOW_LEVEL$ and pension gain has a negative significant effect on retirement, while pension gain by itself is insignificant.

Because we are primarily interested in individuals' knowledge of the financial incentive to retire inherent in their pension structure, we use a third measure of knowledge that captures the accuracy of individuals' pension gain based on comparisons between self- and employer-reports. We form each individual's pension gain using both self- and employer-reported values. Then, the variable $KNOW_GAIN$ is set equal to one if the ratio of pension gain based on self- and employer-reports is between 0.5 and 2, or if one is zero and the other is within \$5,000. Table 1 shows that 31 percent of individuals in our sample are well-informed regarding this more complex measure of the incentive effects of pensions. This measure requires individuals to understand benefit eligibility, amounts and reductions at alternative hypothetical retirement ages.

In column 4 of Table 3, we use $KNOW_GAIN$ and its interaction with pension gain. This specification again produces similar results to the other information measures. The main coefficient on pension gain, capturing the effect of employer reported pension gain for individuals whose self-reports are substantially different than their employer reports (or who do

not report enough information to construct it), is very small and no longer statistically significant. Having reasonably accurate knowledge of one's pension gain significantly increases the probability of retirement by over 2 percentage points, although again, we note that this could reflect the fact that individuals particularly interested in retirement may be more likely to be knowledgeable about their pension plans. Perhaps most interesting, however, is the interaction between *KNOW_GAIN* and pension gain. For those individuals who accurately report the financial incentives implied by their pension structures, there is a very large negative effect of pension gain on retirement probabilities. The coefficient for this group is more than 3 times the size of the average effect for the entire sample (shown in column 1 of the table). The coefficient of implies that a one standard deviation increase in pension gain would lead to a reduction in the retirement probability of 1.8 percentage points, or 27 percent. This highlights some dramatic heterogeneity underlying the average effect of pension gain reported in column 1, and the estimates reported in previous literature. The effect of pension-related financial incentives appears to be driven by a minority of individuals (30 percent of our sample) who are relatively well-informed about their pensions.

Columns 5-7 of Table 3 repeat some of the above analysis using a measure of *GAIN* constructed from both pensions and Social Security. The coefficient on *GAIN* in column 5 is slightly smaller but still statistically significant. The information measures that include individuals' knowledge of their expected Social Security benefits are constructed in a similar fashion to those for private pensions only. Unfortunately, the self-reported data on expected Social Security are much less extensive than the self-reported pension data. Specifically, individuals are asked only for their expected Social Security benefit amount at their expected retirement age. Thus, for Social Security, we can only create a knowledge measure that indicates

whether the expected level of benefits at that expected retirement age is close to the level we calculate from Social Security Administration data, *KNOW_SS*. Specifically, it is not possible to construct a measure similar to *KNOW_GAIN* for Social Security benefits. Another limitation with respect to knowledge of Social Security benefits is that only a single respondent per household was asked to provide expected benefit information. To avoid using a spouse's knowledge of Social Security benefits to determine an individual's behavior, we include only those individuals who directly report their own expected Social Security benefit in this analysis. As shown in column 6 of Table 3, the interaction between knowing one's Social Security benefit level and the incentive to retire (including both pensions and Social Security) is not statistically significant. We view this result primarily as a reflection of the limited data we have available on knowledge of Social Security benefits. In column 7, where we return to using *KNOW_GAIN* as the knowledge measure (knowledge of private pensions only), but instead consider the gain from private pensions and Social Security, we get similar results to that for private pensions alone. The coefficient on *KNOW_GAIN* interacted with Social Security and pension gain is negative and significant, while that on Social Security and pension gain itself is not.

The results in Table 3 are robust to a number of variations in the basic sample. We obtain similar results when we restrict the sample to include only individuals with defined benefit pensions (for whom we have more confidence in the employer-reports as the truth). In addition, eliminating wave 1 from the analysis does not substantially alter the results. We have also tried specifications in which we drop from the sample individuals who do not report enough information to form a self-reported measure of pension gain (*REPORT*=0 and thus, *INFORMED*=0). Looking at the effects of the accuracy of information conditional on having reported all of the key pension components allows us to distinguish between individuals who

simply do not report information on their pensions and those who report values that are very different from their employer-reports. These coefficients are very similar to those reported in Table 3.⁹

We also explored several alternative definitions of the knowledge variables. In particular, we tried altering the definition of *KNOW_GAIN* to allow for different degrees of error in determining who knew their pension gain accurately. For example, we considered definitions in which the ratio of self- to employer-reported pension gain was between 0.75 and 1.25, and other cut-offs. The coefficients on *KNOW_GAIN* were not sensitive to these cutoffs. This suggested that individuals were responding to something other than the precise dollar values of their pension gain. One possibility is that individuals retire when they face a zero gain or a loss from continuing to work. If this is the case, it may be enough if individuals simply understand the sign of their pension gain. To explore this possibility, we constructed one additional measure of pension knowledge, *KNOW_SIGN*, that is equal to one if the employer- and self-reports of pension gain are both positive, both negative, or if both are equal to zero.

The results of interacting *KNOW_SIGN* with pension gain are shown in Table 4. Column 1 simply repeats the first column of Table 3 for ease of comparison. The coefficient on pension

⁹ We have also considered measures that make use of the extent of correlation between the self-reported and employer-based series of pension values. That is, because an individual's ability to predict the pension gain measure properly depends upon their knowledge of differences in pension wealth at different retirement ages, the shape of the retirement-age profile of pension wealth will be critical. Using the series of present values of pension wealth based on retirement from the current age to age 70, we calculate the correlation coefficient between the self- and employer reported series, and include this correlation coefficient, along with its interaction with pension gain, in the retirement regressions. The use of this continuous measure of pension information held by individuals gives mixed results. One difficulty with using this measure is that, for individuals with no variation across retirement ages in their pension profiles (variance of their present values is zero), the correlation coefficient is missing. Additionally, for those who do not report enough pension information to construct the present values, it is less clear for this measure what value to use. In one specification, we set negative correlation values and missing values to zero. This specification produces a negative interaction with pension gain, though the interaction itself is also not statistically significant. The magnitude implies that an individual with a correlation of one between the self- and employer-reported pension wealth profile would have roughly twice the response to pension gain as a similar individual with a zero correlation.

gain here is the average effect of pension gain on retirement. In column 2 of Table 4, we show that there is no statistically significant response to pension gain among those who do not know whether they stand to gain or lose pension wealth from continuing to work (roughly 40% of those self-reporting pension values). In contrast, those who do know the sign of pension gain have the expected negative and significant response to the financial incentive. Moreover, individuals who are aware of the sign of their pension gain have a responsiveness to that pension gain that is approximately equal to that of informed individuals when we use the *KNOW_GAIN* variable (as shown in Table 3, column 4).

Finally, given that the sign of pension gain appears to be an important piece of individual knowledge, we consider whether the sign alone can explain retirement behavior as well as the continuous measure. In column 3 of Table 4 we show results of a model of retirement in which dummies for negative and zero pension gain are included (the omitted category is positive pension gain). The dummy for a negative pension gain is positive and significant, consistent with those who are past their pension wealth peak being substantially more likely to retire; the retirement probability is almost doubled for those with negative pension gain. Surprisingly, the dummy on pension gain equal to zero is negative, although smaller than that for pension gain less than zero. It is also the case that the adjusted R-squared between the first and third columns of the table are virtually identical. When we interact the categories for the sign of pension gain with knowledge of the sign of pension gain, we get the expected results. Individuals who know they have a (positive) gain to continuing work are less likely to retire than those who are unaware of such a gain. Similarly, those aware of zero or negative pension gains are more likely to retire.

V. Interpreting the Interaction between Knowledge and Pension Incentives

The results discussed so far offer important evidence on the potential range of responsiveness to pensions across the population, but they should be interpreted carefully. It is of course possible that some individuals may simply have more information for reasons that are exogenous to their tastes and plans for retirement. For example, some firms may simply be more thorough in communicating pension and other benefit information to their employees. In this case, we could interpret the results above as the causal effect of information provision and pensions on retirement. It seems unlikely, however, that all information acquisition is exogenous to the retirement decision. In this section, we consider two alternative models that address whether providing more or better information to individuals will increase their responsiveness to those incentives.

A. A Retirement Model with Selection into Knowledge States

Individuals may not actively collect information about their pensions until they are seriously considering retirement. If this is the case, unobserved characteristics may make individuals both more likely to be well-informed about their pensions, and more likely to retire. Combining this with a formulation in which there are (potential) differences in the determinants of retirement among the informed and uninformed we have a two stage model that can be written as follows:

$$[3] \quad INFORMED_{it} = \alpha_0 + \beta_1 GAIN_{it} + \beta_2 X_{it} + \beta_3 Z_{it} + v_{it}$$

$$[4] \quad a) \quad R_{it} = \gamma_0^U + \gamma_1^U GAIN_{it} + \gamma_2^U X_{it} + \varepsilon_{it}^U \quad \text{if uninformed } (INFORMED_{it} = 0)$$

$$b) \quad R_{it} = \gamma_0^K + \gamma_1^K GAIN_{it} + \gamma_2^K X_{it} + \varepsilon_{it}^K \quad \text{if informed } (INFORMED_{it} = 1)$$

where $INFORMED_{it}$ is one of several binary variables indicating whether or not an individual is informed about his or her financial options as defined above, X_{it} and Z_{it} are vectors of variables that affect information acquisition, and v is an error term. Note that both $GAIN$ and X will appear in the equation predicting retirement below, but that Z is unique to this first stage.¹⁰ If v_{it} is correlated with the ε_{it} terms, OLS estimation of equations [4] will be subject to a standard selection bias.

As shown by Heckman (1976) and Lee (1976), equations [4] can be consistently estimated by explicitly conditioning on selection into the two regimes (uninformed and knowledgeable). Inverse mills ratios are calculated from the estimation of equation [3] and then included in the estimation of equations [4].

This major challenge in estimating the system of equations [3] and [4] involves the separate identification of the information equation and the retirement equations. To consistently estimate the parameters of the model, we need to identify factors that: (i) help predict the individual knowledge of pension incentives, but that, (ii) do not also affect retirement decisions. Towards this end, we briefly consider what factors help determine an individual's level of pension knowledge.

Previous work provides some guidance as to what factors should determine an individual's level of pension knowledge. Mitchell (1988) proposes a general framework for thinking about information acquisition in this setting, identifying three broad sets of factors that should influence the extent of pension information individuals possess. First, the abilities of individuals to process complex financial information may affect retention and understanding of their pension plans. Educational attainment and other individual characteristics fall into this

¹⁰ We allow for the possibility that $GAIN$ affects the extent of knowledge since models of information acquisition

category. Note that many of these characteristics may also affect retirement probabilities, for example, through correlation with tastes or preferences towards leisure, and so will not help in the separate identification of equations [3] and [4]. Nevertheless, it is important to understand how such characteristics affect knowledge since this will also tell us something about which segments of the population are most likely to respond strongly to pension incentives. One final factor that might influence knowledge is whether an individual is covered by multiple pension plans. This could serve as a proxy for the complexity of the calculations required to summarize the financial incentives associated with alternative retirement dates. We include a dummy variable equal to one if an individual has multiple pension plans. This variable is excluded from the retirement equation, and so contributes to the identification of the selection model.

A second set of factors that should influence pension knowledge involves firm-based differences in the efficiency of information dissemination. Employers have a role to play in providing information to the beneficiaries of their pension programs. Focusing on variables that may lead to information dissemination by pension providers is one way to identify factors that are more likely to be exogenous to workers' retirement behavior. We consider several firm-level variables as determinants of individual's pension knowledge, several of which are excluded from the retirement equation to identify the selection model.

Firm size, which is available in the HRS, is one such factor. Larger firms may be more likely to have formal information programs, or at least human resources departments able to systematically provide information to their workers.¹¹ There is little reason to think that, after conditioning on financial incentives and lifetime earnings, the size of a worker's firm should

suggest that those with more potential benefit will be more likely to invest in information gathering.

¹¹ Mitchell (1988) uses a measure of firm size in her analysis of the determinants of pension information, and does not find it to have a significant effect. She had information, however, only on the average firm size at the industry level while the HRS provides information on the actual firm size of the individual's employer.

influence retirement. Union membership might similarly indicate the existence of an information-provision system for a given worker. Both Mitchell (1988) and Gustman and Steinmeier (2001c) find some evidence that unionized workers are better informed about their retirement incentives. Union status should not influence the retirement decision, as long as retirement incentives and other characteristics of the job that might alter the utility of leisure (physical demands of the job, for example) are controlled for. Fortunately, the HRS also contains a variety of measures of such job characteristics that might allow us to credibly exclude a worker's union status from the retirement regressions.

We propose one other characteristic associated with the firm as potential determinants of an individual's information level. We include a variable indicating the time until a given individual's retirement-wealth maximizing age based on the employer-reported pension data. The idea here is that firms will have an incentive to inform individuals of upcoming pension benefits as these individuals approach important "kinks" in their pension benefit schedules. Conditional on the private pension gain associated with a given year and the individual's age (entered non-linearly), the number of years before or after their optimum should not influence retirement.¹²

A third set of factors identified by Mitchell and others as influencing individual knowledge of pensions involves the benefit of that information. A greater likelihood of benefiting from knowledge of pension details should lead individuals to acquire more information. The size of the pension gain itself maybe relevant here, as well as health status, job characteristics and the age and health of the individual and his or her spouse. As was the case

¹² This is similar in spirit to Luchak and Gunderson's (1999) finding that individuals at or beyond their wealth maximizing age are more knowledgeable about pension benefits.

with individual demographic characteristics, most of these factors will also affect the retirement decision.

Based on the above considerations, we estimate the model summarized by equations [3] and [4] using time left until the pension maximizing age from employer reports, and indicators for multiple pension plans, firm size, and union status as the excluded variables in estimating the selection model.

Results from stage 1: predicting INFORMED

Consistent with our earlier findings, we show results for three separate knowledge measures: *REPORT*, *KNOW_GAIN*, and *KNOW_SIGN* in our estimate of [3]. While the latter two measures have the strongest effects on individuals' responsiveness to their pension gain, the first is also of interest since it will show the extent of non-random selection that occurs when using the self-reported pension data.

Results from probit models for these knowledge measures are shown in Table 5. Several factors have common effects on all of the knowledge measures. Black and Hispanic individuals are less well-informed, regardless of the information measure used. The probability of knowing the sign of pension gain among blacks, for example, is 10 percentage points lower than that for comparable white individuals; for Hispanics the reduction is approximately 6 percentage points.¹³

The large effects of race and ethnicity on information provide an opportunity to demonstrate the importance of understanding differential responsiveness among informed and uninformed individuals, we compare retirement behavior by race. In Table 6, we repeat

¹³ These and subsequent effects cited are calculated at the sample means of continuous variables, and for a white, male, non-union worker at a small firm with a high school education, unmarried and without multiple pension plans.

estimates of the basic retirement equations that were shown in Table 3, splitting the sample by race. In column 1, when we consider pension gain only, we find no significant effect of pension gain on the probability of retirement for Blacks and Hispanics, in contrast to the significant result for Whites in column 2. These average population results would lead us to conclude that Blacks and Hispanics are unresponsive to pension incentives, while Whites are responsive. However, in columns 3 and 4 where we account for knowledge, we find that Blacks and Hispanics who are informed have a significant response to pension gain, and that this response is even larger than the response among informed Whites. For these subpopulations, the difference between the results when we do and do not account for knowledge is striking.

Education does not seem to significantly increase the probability of being well-informed, and actually decreases the probability of having a self-reported pension gain that is close in magnitude to that based on the employer reports (*KNOW_GAIN*=1). Health and several job characteristics (whether the job is physically demanding, stressful, or difficult) also do not have consistent or significant effects on knowledge. There is suggestive evidence that men are better informed, although this does not hold for the *KNOW_GAIN* measure.

We have also included dummy variables for survey waves, and for individuals' age. Somewhat surprisingly, knowledge seems to peak at wave 2, and then declines slightly. This decline may also reflect the fact that the employer reports are less likely to capture the "true" pension since they are measured at the time of the wave 1 survey.¹⁴ There is some evidence of improved information as individuals age when *REPORT* and *KNOW_GAIN* are the knowledge measures. The difference in the probability of being well-informed about one's pension gain

¹⁴ Recall that job changers and those who report changes to their pensions in later waves have been dropped from our samples.

between age 55 and age 60 is approximately 4 percentage points. When *KNOW_SIGN* is our measure of knowledge, however, there is evidence that knowledge actually declines with age.

The evidence from Table 5 also suggests that the size of both pension gain and the present value of pension wealth influence the likelihood of being well informed, although this varies somewhat depending upon the measure used. For both *KNOW_GAIN* and *KNOW_SIGN*, having a larger gain to deferring retirement makes an individual more likely to be well-informed, consistent with theories suggesting that those with larger marginal benefits of information are more likely to gather that information.

The variables at the top of Table 5 are those that we use to predict the knowledge measures but exclude from the retirement equation. As expected, individuals with multiple pension plans are significantly less likely to report all of the necessary pension details, or to report them accurately. Having multiple pension plans reduces the probability of being well-informed by 2 to 5 percentage points, depending upon the knowledge measure used. Union members are 5 percentage points more likely to fully report pension information, and 7 percentage points more likely to be aware of whether they have a gain, a loss, or no change in retirement wealth from deferring retirement. Surprisingly, the union variable is significant and negative in the equations predicting *KNOW_GAIN*. Union members are roughly 4 percentage points less likely to precisely identify the magnitude of their pension gain. The firm size measures also have unexpected signs and are generally not statistically significant. Where they are significant, they suggest, contrary to our expectations, that those working for firms of greater than 100 employees are less well-informed about their pensions.

Finally, the variable indicating years until an individual's pension-wealth-maximizing age is negative (as expected) and significant in two of the three columns of Table 5. Being one

year closer to the wealth-maximizing retirement age makes individuals roughly half a percentage point less more likely to come close to knowing their pension gain and about one percentage point more likely to know whether their pension gain is positive, zero or negative.

Results from stage 2: the effects of information, controlling for selection

The results of estimating equation [4] including the selection correction terms are summarized in Table 7. Selection into the informed and uninformed regime is predicted using the specification summarized in Table 5. The first two columns of Table 7 show a model of retirement that allows the variable *REPORT* to be non-random. The terms for pension gain and pension wealth in this model are very imprecisely estimated, and none are statistically significant. We also cannot reject the hypothesis that the coefficients on the pension terms in the informed (*REPORT*=1) and uninformed (*REPORT*=0) regimes are equal to one another. The selection terms are statistically significant, suggesting that fully reporting key pension details is a significant determinant of retirement. At the bottom of the table, we also include the comparable coefficients from OLS retirement regressions estimated separately for the informed and uninformed groups. Here, the coefficients on the pension terms are again similar to one another, but are more precisely estimated than in the two stage model.

The next two columns of Table 7 repeat this exercise using the variable *KNOW_GAIN* as the indicator of knowledge. As expected given our earlier results, there are very different effects of pension gain and pension wealth among the informed and uninformed groups. Among those with *KNOW_GAIN*=1, a one-standard deviation increase in pension gain (\$53,000) reduces the probability of retirement by 2 percentage points, or approximately 30 percent. Among those with *KNOW_GAIN*=0, there is not a significant effect of pension gain on the probability of

retirement. The present value of pension wealth is positive and significant in both regimes, but the magnitude of the effect is more than twice as large among the more informed segment of the population. As before, the selection terms are statistically significant, indicating non-random selection into the informed state. Finally, the OLS coefficients on the pension terms (when the regressions are estimated separately by knowledge status) are very similar to those estimated based on the selection model.

The results in the columns five and six of Table 7 echo the results based on *KNOW_GAIN*. When *KNOW_SIGN* is used as our measure of pension knowledge, there are large differences in the extent of responsiveness to pension gain (but not pension wealth levels in this case), and this finding remains whether or not we control for selection into the regimes. The selection correction term is statistically significant in the informed regime, but not in the uninformed state.

Thus far, the selection correction model shows that there does appear to be significant correlation across the equations determining knowledge and retirement behavior, but explicitly controlling for this correlation does not appear to drive differences in the responsiveness of individuals to their pension incentives. It appears that there is a fairly weak correlation between pension gain and the likelihood of becoming informed.

In columns 7 and 8 of Table 7 we estimate a slightly different specification of the selection model, that includes both a continuous pension gain variable, and a dummy for having pension gain less than or equal to zero. This specification is motivated by our concern that individuals with non-positive pension gains appear both more likely to retire, and much more likely to be well-informed. If the relationship between pension gain and the likelihood of becoming informed is highly non-linear, this could explain why the sample-selection correction

makes little difference to our results. The results in the final two columns do show that having pension gain less than or equal to zero is a strong predictor of both retirement and the extent of pension knowledge. In this case, the selection correction does change the coefficients on pension gain and pension gain less than zero, but qualitatively, our basic results remain. There is little or no effect of pension incentives among the uninformed group, and significant effects of these incentives among those individuals who are well-informed.

B. A Retirement Model with Endogenous Information Acquisition

Another possible explanation for our finding of stronger response among the knowledgeable is that individuals might determine how much information to gather based on how likely they are to take account of that information in making retirement decisions. Suppose that there is a distribution of responsiveness to financial incentives across the population, and that this individual-specific responsiveness is a key factor in determining the distribution of pension knowledge across the population. In such a model where heterogeneous responses to pension incentives drives heterogeneity in pension knowledge, our basic results, again, cannot be interpreted as the effect of knowledge on the responsiveness to pension incentives. This poses a more difficult challenge for estimation of causal effects, and is similar to the issues raised in the literature on estimating heterogeneous returns to education (see Card, 1999, for a discussion). Given the lack of any convincing instruments for pension information in this situation, we cannot formally rule out the possibility that causality runs from dispersion in responsiveness to dispersion in information.

We do however, provide some evidence against this possibility by returning to the role that the self-reported pension gain itself plays in retirement decisions. To simplify the

hypothesis described above, suppose there two groups of individuals in the population: those who respond to financial incentives in making their retirement decision and those who do not. Only the first group would bother to collect pension information, and the second group's behavior would not be altered, even if information was exogenously provided to them. Under such a scenario, this second group should not respond to the true pension information (the employer-reports) but should also not respond to their self-reported information. They may have a best guess of their pension knowledge, but if they are simply "non-responders", we should not see them responding to that information. Table 8 produces results analogous to those in Table 3, but using the self-reported pension information (rather than employer-reports) interacted with the knowledge measures. In the first column we see that individuals whose self-reported pension gain are not particularly close to their true pension gain (the coefficient on pension gain by itself) nevertheless respond to that incentive in making their retirement decisions. The better-informed group shows an even larger response to self-reported pension gain, as shown by the interaction of *KNOW_GAIN* and pension gain. The results in the second column are even more striking: there is no difference in the responsiveness of individuals to their own perceptions of pension gain based on the accuracy of those self-reports, when our measure of accuracy is simply whether they are aware of the sign of their incentive. This is inconsistent with a model in which the uninformed are simply those who would not respond to pension information under any circumstances.

VI. Conclusion

This paper makes two important contributions to the large literature exploring the link between financial incentives embedded in private pensions, and individual retirement decisions. First, we show that there is substantial heterogeneity in whether individuals take account of financial factors in choosing a retirement date, and that this heterogeneity can be directly related to individuals' level of knowledge about these financial factors. Average effects of pensions on retirement reflect a combination of effects among well-informed segments of the population that are two to three times larger than this average, and effects among the ill-informed that are essentially zero. To understand or predict the retirement response to changes in pension or Social Security policies, it may be important to consider the knowledge level of intended beneficiaries of those policies.

A second contribution of this work is to point to the usefulness of self-reported pension data, as collected in the Health and Retirement Study and other surveys. Individuals do respond directly to these self-reports, suggesting that they do capture the pension incentives that individuals perceive themselves to be facing. The self-reported pension data provide an important intermediate step between administrative reports of pensions and the response of individual retirement behavior to these pensions.

Lastly, this paper has important implications beyond understanding retirement behavior. Much applied economic research seeks to measure the responsiveness of individuals to certain incentives. At a minimum, our results point to the need to think carefully about whether those incentives are properly known by the individuals that are assumed to be affected.

References

- Card, David (1999) "The causal effect of education on earnings" in *Handbook of Labor Economics* Volume 3, Orley Ashenfelter and David Card, eds.
- Chan, Sewin, and Ann Huff Stevens (2002) "How Does Job Loss Affect the Timing of Retirement?" NBER Working Paper 8780.
- Chan, Sewin, and Ann Huff Stevens (2003) "Do Changes in Pension Incentives Affect Retirement? A Longitudinal Study of Subjective Retirement Expectations" *Journal of Public Economics*, forthcoming.
- Choi, James, et al. (2001) "Defined Contribution Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance." NBER Working Paper #8655.
- Coile, Courtney and Jonathan Gruber (2000) "Social Security and Retirement," NBER Working Paper 7830.
- Duflo, Esther and Emmanuel Saez (2002) "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices" *Journal of Public Economics* 85(1): 121-48.
- Gustman, Alan L., and Thomas L. Steinmeier (2001a) "What People Don't Know about their Pensions and Social Security" in William G. Gale, John B. Shoven and Mark J. Warshawsky, editors, *Public Policies and Private Pensions*, Washington D.C.: Brookings Institution. (NBER Working Paper 7368.)
- Gustman, Alan L., and Thomas L. Steinmeier (2001b) "Retirement and Wealth," *Social Security Bulletin*, Perspectives, Vol. 65 No.3, forthcoming. (NBER Working Paper 8229.)
- Gustman, Alan L., and Thomas L. Steinmeier (2001c) "Imperfect Knowledge, Retirement and Saving," NBER Working Paper 8406.
- Heckman, James J. (1976) "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5 Fall: 475-92.
- Lee, Lung Fei. "Estimation of Limited Dependent Variables Models by Two-Stage Methods." Ph.D. Dissertation, University of Rochester, 1976.
- Luchak, Andrew, and Morley Gunderson (2000) "What Do Employees Know About their Pension Plan?" *Industrial Relations* 39:4, 646-670.

- Lumsdaine, Robin, James Stock and David Wise (1992) “Three Models of Retirement: Computational Complexity Versus Predictive Validity,” in *Topics in the Economics of Aging*, David Wise, ed. Chicago: University of Chicago Press.
- Mitchell, Olivia (1988) “Worker Knowledge of Pension Provisions,” *Journal of Labor Economics* 6:1, 21-39.
- Samwick, Andrew (1998) “New Evidence on Pensions, Social Security, and the Timing of Retirement,” *Journal of Public Economics* 70:2, 207-36.
- Stock, James and David Wise (1990) “Pensions, the Option Value of Work, and Retirement,” *Econometrica* 58: 1151-1180.
- Uccello, Cori, and Kevin Perese (1999) “Wealth Accumulation in the Health and Retirement Study: The Importance of Including Pension Wealth,” Urban Institute Report.

Table 1
Sample Summary Statistics

	Mean	Standard deviation
% Retired this year	0.067	
% Retired by end of sample	0.34	
% Male	0.54	
% Hispanic	0.05	
% Black	0.16	
Age	57.2	3.7
% Married	0.79	
Total non-pension assets \$	249,023	590,789
Average Lifetime Earnings \$	22,894	11,607
Pension wealth if retire immediately (self-report) \$	110,881	310,133
Pension wealth if retire immediately (employer-report) \$	147,558	245,054
Pension gain (self-report) \$	84,393	177,576
Pension gain (employer-report) \$	31,462	52,928
% Reporting key values of pension (<i>REPORT</i> =1)	0.60	
% Self-reporting expected SS benefits	0.37	
Among those reporting:		
% With accurate level of pension wealth (<i>KNOW_LEVEL</i> =1)	0.50	
% With accurate pension gain (<i>KNOW_GAIN</i> =1)	0.31	
% With accurate SS level (<i>KNOW_SS</i> =1)	0.78	
% With knowledge of sign of pension gain (<i>KNOW_SIGN</i> =1)	0.60	
Note: accurate implies the ratio of employer to self-report is between 0.5 and 2		
No. of observations	15,930	

Notes: Pooled data from waves 1-4 of the HRS. Sample is described in section II of the text.

Table 2
The Effect of Self- and Employer-Reported Pension Gain on the Annual Probability of Retirement

Dependent Variable:	Retire				
Estimation procedure:	OLS 1	OLS 2	OLS 3	OLS 4	IV 5
Pension gain based on employer-report	-0.010 (0.003)	-0.012 (0.005)		-0.007 (0.005)	
Pension wealth level based on employer-report	0.005 (0.002)	0.005 (0.001)		0.006 (0.002)	
Pension gain based on self-report			-0.008 (0.001)	-0.010 (0.002)	-0.021 (0.012)
Pension wealth level based on self-report			0.003 (0.002)	0.002 (0.002)	0.040 (0.011)
Number of observations	14,288	8,600	8,600	8,600	8,600
Corr(pension gain based on employer-report, pension gain based on self-report)				0.22	
Corr(pension wealth level based on employer-report, pension wealth level based on self-report)				0.21	

Notes: Estimated using data from waves 1-4 of the HRS. Standard errors are in ().

Pension gain and wealth levels are in \$100,000s.

Other explanatory variables include race, education, health, marital status, non-pension assets,

Table 3
The Effect of Pension Gain and Knowledge on the Annual Probability of Retirement

Dependent Variable: Estimation procedure:	Retire OLS						
	1	2	3	4	5	6	7
Pension gain	-0.010 (0.004)	0.004 (0.005)	-0.004 (0.004)	-0.002 (0.004)			
<i>REPORT</i> =1		0.023 (0.005)					
(<i>REPORT</i> =1)*Pension gain		-0.022 (0.006)					
<i>KNOW_LEVEL</i> =1			0.016 (0.005)				
(<i>KNOW_LEVEL</i> =1)*Pension gain			-0.015 (0.007)				
<i>KNOW_GAIN</i> =1				0.025 (0.007)			0.034 (0.010)
(<i>KNOW_GAIN</i> =1)*Pension gain				-0.034 (0.007)			
Pension and Social Security gain					-0.007 (0.003)	-0.008 (0.008)	-0.004 (0.003)
<i>KNOW_SS</i> =1						-0.001 (0.014)	
(<i>KNOW_SS</i> =1)*Pension and Social Security gain						-0.001 (0.008)	
(<i>KNOW_GAIN</i> =1)*Pension and Social Security gain							-0.016 (0.006)
Number of observations	14,288	14,288	14,288	14,288	14,287	5,851	14,287
R-squared	0.24	0.24	0.24	0.24	0.24	0.22	0.24

Notes: OLS coefficients estimated using data from waves 1-4 of the HRS. Standard errors are in (). Pension gain and wealth levels are in \$100,000s and are based on employer-reports. Other explanatory variables include race, education, health, marital status, non-pension assets, age, gender, wave, retiree health insurance, lifetime earnings and an intercept.

Table 4
The Effect of Knowing the Sign of Pension Gain on the Annual Probability of Retirement

Dependent Variable:	Retire			
Estimation procedure:	OLS			
	1	2	3	4
Pension gain	-0.010 (0.004)	0.004 (0.005)		
<i>KNOW_SIGN</i> =1		0.012 (0.005)		-0.023 (0.003)
<i>(KNOW_SIGN</i> =1)*Pension gain		-0.031 (0.006)		
Pension gain=0			-0.021 (0.005)	-0.045 (0.006)
Pension gain<0			0.063 (0.008)	0.024 (0.009)
<i>(KNOW_SIGN</i> =1) and (Pension gain=0)				0.068 (0.012)
<i>(KNOW_SIGN</i> =1) and (Pension gain<0)				0.144 (0.019)
Number of observations	14,288	14,288	14,288	14,288
Adjusted R-squared	0.24	0.24	0.25	0.26

Notes: OLS coefficients estimated using data from waves 1-4 of the HRS. Standard errors are in (). Pension gain and wealth levels are in \$100,000s and are based on employer-reports. Other explanatory variables include race, education, health, marital status, non-pension assets, age, gender, wave, retiree health insurance, lifetime earnings and an intercept.

Table 5
Determinants of Pension Knowledge

Dependent Variable	<i>REPORT</i>	<i>KNOW_GAIN</i>	<i>KNOW_SIGN</i>
Estimation procedure:		Probit	
	1	2	3
Excluded from retirement equation:			
Multiple pension plans	-0.127 (0.052)	-0.121 (0.051)	-0.087 (0.046)
Years till maximum pension wealth	0.007 (0.007)	-0.021 (0.006)	-0.021 (0.005)
Union member	0.128 (0.049)	-0.184 (0.051)	0.182 (0.043)
Firm size > 100 and < 500 employees	-0.103 (0.055)	-0.055 (0.056)	-0.065 (0.050)
Firm size > 500 employees	-0.007 (0.055)	-0.201 (0.058)	-0.026 (0.049)
Firm size missing	-0.291 (0.153)	0.492 (0.153)	0.249 (0.151)
Included in retirement equation			
Hispanic	-0.235 (0.106)	-0.176 (0.113)	-0.158 (0.100)
Black	-0.336 (0.063)	-0.255 (0.072)	-0.286 (0.058)
Some college	0.036 (0.060)	0.060 (0.062)	0.021 (0.054)
College graduate	0.027 (0.069)	-0.163 (0.061)	-0.001 (0.055)
Couple	0.033 (0.055)	0.027 (0.062)	0.055 (0.052)
Non-pension assets	-0.004 (0.003)	-0.008 (0.005)	-0.008 (0.004)
Average lifetime annual earnings	0.501 (0.245)	0.430 (0.260)	0.054 (0.213)
Average lifetime earnings missing	-0.117 (0.080)	0.004 (0.088)	-0.130 (0.074)
Pension gain	-0.070 (0.052)	0.156 (0.045)	0.217 (0.053)
Pension wealth level	0.098 (0.034)	-0.003 (0.009)	0.032 (0.012)
Bad health	-0.022 (0.065)	0.017 (0.066)	0.020 (0.061)
Job involves physical effort	-0.063 (0.030)	-0.012 (0.032)	-0.017 (0.028)
Job is more difficult than before	0.045 (0.032)	-0.045 (0.034)	0.065 (0.029)
Job involves stress	-0.047 (0.033)	-0.065 (0.032)	-0.056 (0.029)
Male	0.089 (0.056)	-0.016 (0.057)	0.094 (0.048)
Wave 2	0.354 (0.039)	0.066 (0.043)	0.125 (0.037)
wave 3	0.219 (0.049)	-0.088 (0.054)	-0.044 (0.046)
Wave 4	0.078 (0.091)	-0.101 (0.103)	-0.279 (0.096)

Age 52	0.006 (0.046)	-0.066 (0.059)	-0.033 (0.049)
Age 53	0.021 (0.059)	0.004 (0.071)	0.019 (0.060)
Age 54	0.009 (0.067)	-0.055 (0.077)	0.006 (0.067)
Age 55	0.017 (0.069)	-0.179 (0.081)	-0.150 (0.071)
Age 56	0.036 (0.074)	-0.160 (0.085)	-0.121 (0.074)
Age 57	0.070 (0.078)	-0.186 (0.089)	-0.110 (0.077)
Age 58	0.038 (0.080)	-0.149 (0.090)	-0.100 (0.080)
Age 59	0.063 (0.083)	-0.120 (0.093)	-0.119 (0.082)
Age 60	0.055 (0.087)	-0.169 (0.097)	-0.205 (0.085)
Age 61	0.034 (0.092)	-0.096 (0.102)	-0.176 (0.090)
Age 62	0.034 (0.098)	0.007 (0.106)	-0.282 (0.095)
Age 63	-0.018 (0.107)	0.060 (0.117)	-0.309 (0.105)
Age 64	-0.005 (0.121)	0.206 (0.129)	-0.263 (0.118)
Age 65	0.048 (0.135)	0.167 (0.144)	-0.477 (0.137)
Age 66	0.003 (0.164)	0.251 (0.172)	-0.332 (0.164)
Age 67	-0.123 (0.211)	0.274 (0.222)	-0.289 (0.217)
Age 68	0.166 (0.285)	0.522 (0.277)	-0.309 (0.277)
Age 69	-0.162 (0.367)	0.560 (0.378)	-0.338 (0.377)
Intercept	-0.056 (0.129)	-0.624 (0.133)	-0.319 (0.116)
Number of observations	12,620	12,620	12,620

Notes: Probit coefficients estimated using data from waves 1-4 of the HRS. Standard errors are in (). Pension gain and wealth levels are in \$'000s and are based on employer-reports.

Table 6
The Effect of Pension Gain and Knowledge on the Annual Probability of Retirement by Race

Dependent Variable:	Retire			
Estimation procedure:	OLS			
Sample:	Non-white		White	
	1	2	3	4
Pension Gain	-0.005 (0.011)	0.008 (0.015)	-0.010 (0.004)	-0.003 (0.004)
<i>KNOW_GAIN</i> =1		0.035 (0.016)		0.023 (0.008)
<i>(KNOW_GAIN</i> =1)*Pension gain		-0.040 (0.013)		-0.034 (0.008)
Number of observations	2,908		11,434	

Notes: OLS coefficients estimated using data from waves 1-4 of the HRS. Standard errors are in (). Pension gain and wealth levels are in \$100,000s and are based on employer-reports.

Non-white is Black and Hispanic only.

Other explanatory variables include education, health, marital status, non-pension assets, age, gender, wave, lifetime earnings and an intercept.

Table 7
The Effect of Pensions on Retirement with Selection into Knowledge States

Dependent Variable: Measure of <i>INFORMED</i> : Value of <i>INFORMED</i> :	Retirement							
	<i>REPORT</i>		<i>KNOW_GAIN</i>		<i>KNOW_SIGN</i>		<i>KNOW_SIGN</i>	
	=1 1	=0 2	=1 3	=0 4	=1 5	=0 6	=1 7	=0 8
Selection model results								
Pension Gain	-0.002 (0.005)	-0.006 (0.005)	-0.037 (0.009)	0.003 (0.004)	-0.032 (0.006)	0.002 (0.005)	0.0003 (0.006)	-0.003 (0.005)
Pension Gain <=0							0.083 (0.008)	-0.008 (0.006)
Pension Wealth Level	-0.001 (0.002)	0.002 (0.003)	0.011 (0.002)	0.004 (0.000)	0.005 (0.001)	0.005 (0.001)	0.006 (0.001)	0.005 (0.001)
Selection term	-0.156 (0.036)	0.078 (0.038)	0.057 (0.029)	-0.151 (0.026)	-0.072 (0.027)	0.003 (0.026)	0.039 (0.020)	0.018 (0.021)
OLS results								
Pension Gain	-0.010 (0.005)	-0.010 (0.004)	-0.039 (0.008)	-0.002 (0.004)	-0.025 (0.005)	0.003 (0.005)	-0.002 (0.004)	-0.001 (0.006)
Pension Gain <=0							0.088 (0.012)	-0.090 (0.071)
Pension Wealth Level	0.004 (0.001)	0.008 (0.004)	0.011 (0.005)	0.004 (0.002)	0.006 (0.002)	0.005 (0.002)	0.005 (0.002)	0.006 (0.002)
Number of observations	12,626		12,626		12,626		12,626	

Notes: Selection model estimated using data from waves 1-4 of the HRS. Standard errors are in (). Pension gain and wealth levels are in \$100,000s and are based on employer-reports.

Table 8
The Effect of Self-Reported Pension Gain and Knowledge on the Annual Probability of Retirement

Dependent Variable:	Retire	
Estimation procedure:	OLS	
	1	2
Pension gain (self-reported)	-0.007 (0.001)	-0.008 (0.002)
<i>KNOW_GAIN</i> =1	0.004 (0.007)	
<i>(KNOW_GAIN=1)*Pension gain</i>	-0.015 (0.005)	
<i>KNOW_SIGN</i> =1		-0.005 (0.006)
<i>(KNOW_SIGN=1)*Pension gain</i>		0.001 (0.003)
Number of observations	8,600	8,600

Notes: OLS coefficients estimated using data from waves 1-4 of the HRS. Standard errors are in (). Pension gain is in \$100,000s and is based on self-reports. Other explanatory variables include race, education, health, marital status, non-pension assets, age, gender, wave, lifetime earnings and an intercept.