Labor Supply Responses to the Social Security Tax-Benefit Link[*]

Jeffrey B. Liebman

Erzo F.P. Luttmer

David G. Seif

December 22, 2006

Abstract

A key question for Social Security reform is whether workers currently perceive the link on the margin between the Social Security taxes they pay and the Social Security benefits they will receive. We develop a methodology for estimating the incentive effects of the marginal Social Security benefits that accrue with additional earnings on three measures of labor supply: hours, labor earnings, and retirement. We show how three sources of discontinuities in the marginal Social Security benefit formula can be exploited to identify these incentive effects. We find that the retirement hazard rate starts to increase sharply when the current year's earnings crowd out a prior year's earnings in the Social Security benefit formula. This result is consistent with individuals responding to incentives implicit in the Social Security benefit formula, but further analysis is needed to determine whether the Social Security rules cause this result.

[*] Liebman and Luttmer: Kennedy School of Government, Harvard University and NBER. Seif: Economics Department, Harvard University. Corresponding author: Erzo Luttmer, erzo_luttmer@harvard.edu. This research was supported by the U.S. Social Security Administration through grant #10-P-98363-1-03 to the National Bureau of Economic Research as part of the SSA Retirement Research Consortium. The findings and conclusions expressed are solely those of the authors and do not represent the views of SSA, any agency of the Federal Government, or the NBER. All errors are our own.

1. Introduction

It is commonly argued that investment-based Social Security reform will improve economic efficiency by increasing the perceived link between retirement contributions and retirement benefits (Kotlikoff, 1996; Feldstein and Liebman, 2002). Under this argument, individuals perceive the OASDI payroll tax as a pure tax – they fail to recognize that the payment of Social Security taxes will increase their future Social Security benefits. With personal retirement accounts, in contrast, the link between contributions and future income would be clear, and the economic distortions would be reduced.

There is little evidence, however, as to whether people perceive the Social Security tax as a pure tax or whether they realize that the *effective* marginal Social Security tax rate (the nominal tax rate minus the marginal Social Security benefit rate) is generally lower than the nominal Social Security tax rate. To our knowledge, no papers have examined whether the marginal Social Security benefit rate affects labor supply as measured by hours or earnings. While there is an extensive literature on the effect of Social Security on retirement, Diamond and Gruber (1999) note that most of this literature ignores the effect of the marginal Social Security benefit rate (focusing instead on the wealth effects) and that there has been little serious effort to check that the variation in Social Security benefits is uncorrelated with unobserved determinants of retirement.[1] Recent work by Coile and Gruber (2000) and Gustman and Steinmeier (2005) provides strong evidence that retirement decisions do respond to the way in which Social Security benefits vary with retirement dates. However, the rules regarding how benefits vary with early or delayed claiming may be more salient than other provisions determining benefit levels. It is, therefore, unclear whether this evidence is applicable to other potential incentive effects of the Social Security benefit formula. Thus, the purpose of this paper is to develop an empirical methodology that allows us to estimate the extent to which people respond to the marginal Social Security benefits they receive when they increase their earnings.

[1] Notable exceptions are Krueger and Pischke (1992) and Samwick (1998). Krueger and Pischke use the natural variation generated by the "notch generation," but the effect of the marginal Social Security benefit rate is relatively imprecisely estimated, possibly because they use cohort-level data. Samwick shows that retirement decisions respond to discontinuities in marginal retirement benefits associated with reaching the early and full-benefit retirement ages in private defined benefit plans.

The Social Security benefit formula contains a number of discrete jumps and non-linearities that create variation in the effective marginal benefit rate for otherwise very similar individuals (Feldstein and Samwick, 1992; Goda, Shoven, and Slavov, 2006). In particular, we show how variation from the following three features of the Social Security benefit formula can be exploited: (i) Social Security benefits depend only on the 35 highest years of indexed earnings, thus creating jumps in effective Social Security tax rates that depend on which years are included among the 35 highest years. (ii) The 1993 Omnibus Budget Reconciliation Act increased the fraction of Social Security Benefits that would be added to taxable income for higher-income individuals, thus increasing effective Social Security tax rates for some individuals. (iii) The rules for the Social Security benefits of widows and widowers create discontinuities in both the total level of Social Security benefits and marginal Social Security benefits. These discontinuities vary by the earnings of the individual relative to those of the deceased spouse. We show how these three sources of variation can be exploited to estimate whether the effective marginal Social Security tax rate affects labor supply on the extensive margin, as measured by retirement, as well as on the intensive margin, as measured by hours and earnings.

We perform our estimation using observations from the original cohort of the Health and Retirement Study (HRS). The HRS is a longitudinal survey of individuals born between 1931 and 1941 as well as their spouses. The original cohort has been interviewed every two years starting in 1992. We obtained permission to link HRS observations to their administrative Social Security records.

This paper contains preliminary results on the effect of the marginal Social Security benefit rate on retirement decisions using the variation in retirement incentives resulting from the fact that only the 35 highest years of indexed earnings enter into the Social Security benefit formula. In future research, we will also exploit the other two sources of variation, namely variation due to the 1993 tax reform act and due to the Social Security rules for the benefits of widows and widowers. We find mixed evidence on the labor supply response to the retirement incentives created by Social Security. Basic graphs show that the retirement hazard rate is approximately flat in years when current earnings do not replace previous earnings in the Social Security benefit formula, but starts to rise sharply exactly when current earnings start to replace prior years' earnings. However, in our regression models, which include a very rich set of controls (including a full set of age dummies), the evidence on the incentive effect is less pronounced. In some specifications, we find that when marginal Social Security benefits are lower, individuals are significantly more likely to retire. However, in other specifications, this effect is not statistically significant, though we never find a statistically significant negative effect. In future research, we will explore additional specifications and determine which specifications are most credible.

The rest of this paper proceeds as follows. In section 2, we review the relevant literature. Section 3 explains our identification strategy, the data, and our empirical specifications. Section 4 presents the results, and section 5 concludes.

2. Related Literature

There has been a great deal of literature dealing with the incentives that Social Security creates. Kotlikoff (1996) argues that the privatization of Social Security can produce substantial welfare gains by increasing the Social Security tax-benefit linkage. Feldstein and Samwick (1992) calculate the effective Social Security marginal tax rates for different types of individuals and find that these rates vary widely from negative values (for older males with dependent spouses) to positive 10.6 percent (for people who gain no additonal benefits from incremental work). Other papers have focused on how to better link Social Security taxes and benefits, aiming to avoid, for example, the high effective taxes that Feldstein and Samwick find, as well as to increase labor force participation rates of older individuals. Goda, Shoven, and Slavov (2006) and Laitner and Silverman (2006) discuss reforms to the Social Security system that would increase incentives for older workers to remain in the labor force.

Diamond and Gruber (1999) and Feldstein and Liebman (2002) point out that much of the empirical literature on the effects of Social Security has focused on wealth effects rather than on marginal incentives. However, several papers have incorporated the marginal incentives from Social Security into structural models of retirement. For example, in Gustman and Steinmeier's (2005) model, the spike in retirement at age 65 occurs because adjustments to Social Security benefits for those who delay claiming beyond 65 were not actuarially fair in the time period they studied.

Other papers, building upon the Stock and Wise (1990) option value framework, have estimated reduced form models of retirement that suggest that Social Security benefit accrural affects retirement. For example, Coile and Gruber (2000) find that forward looking measures of Social Security benefit incentives are strong predictors of retirement behavior. Samwick (1998) finds that the rate of retirement wealth accumulation (including both private pensions and Social Security) is a significant determinant of the retirement decision, with individuals who are acquiring retirement wealth more rapidly being less likely to retire. In addition, Samwick finds that a high option value of accrual of retirement wealth also makes individuals less likely to retire.

Another important paper that suggests that retirement decisions reflect Social Security benefit incentives is Krueger and Pischke (1992). Although the main emphasis of the Krueger and Pischke study is wealth effects, they also find that continued Social Security wealth accrual lowers the retirement probability.

While these studies indicate that people on the verge of retirement are aware of some of the incentives from the Social Security benefit formula, they do not necessarily imply that people perceive and respond to Social Security incentives more broadly. The particular rules regarding how benefits are adjusted for early or delayed claiming may be more salient than other aspects of the Social Security benefit formula. Thus, it will be important for us to investigate whether people are equally responsive to other sources of variation in benefits.

3. Data and Metholodology

3.1 Discontinuities in the Social Security Benefit Schedule

The complex Social Security rules lead to a large number of discontinuities in Social Security benefits. We identified ten discontinuities, but some of these depend on variables not recorded in our data set, apply to relatively few individuals, or are hard to exploit because they depend on past, and possibly endogenous, choices.[2] This leaves us with the following three sources of discontinuities.

First, we exploit the fact that Social Security benefits depend on only the 35 highest years of indexed earnings (the "maximum-35-year rule"). Thus, after 35 years of earnings, an additional year of earnings will increase benefits only inasmuch as the additional year of earnings exceeds a year of lower earnings. This implies that the effective marginal Social Security benefits from an additional year of work are lower for an individual the higher are the earnings in the year replaced. This individual would therefore have a stronger incentive to retire. We estimate the effect of this incentive by comparing retirement decisions of individuals who are similar in most respects (lifetime earnings, recent earnings, household assets, Social Security wealth, age, spousal age, and other demographic characteristics) except for the level of earnings displaced by an additional year of work. If the earnings in the additional year of work exceed the lowest earnings among the 35 highest years of earnings (as is typically the case), then the *average* returns from working the additional year are reduced. However, the *marginal* returns to working an additional hour in the additional year are unaffected because, on the margin, additional earnings do not displace prior earnings. Thus, the variation in timing of earnings affects incentives to retire, but not labor supply incentives along the intensive margin. Hence, we will use this source of variation only to examine retirement decisions.

Second, the 1993 Omnibus Budget Reconciliation Act increased the fraction of Security Benefits that would be added to taxable income from 50% to 85% for

[2] Other sources of discontinuities in the Social Security schedule include: (1) Treatment of spousal benefits – two individuals with identical earnings histories may face difference marginal Social Security benefits depending on whether it is optimal to claim one's own Social Security benefits or spousal Social Security benefits. (2) Divorce – eligibility for spousal benefits only starts after 10 years of marriage, thus creating a discontinuity in marginal Social Security benefits for individuals who are planning to claim spousal benefits. (3) Remarriage – individuals lose eligibility for spousal benefits based on an ex-spouse upon remarriage, thus creating jumps in marginal Social Security benefits upon remarriage for the subgroup of individuals who would have claimed benefits based on a previous spouse's earnings. (4) The Windfall Elimination Provision – this provision places workers who receive a pension from work and who did not fall under Social Security coverage on a different benefit schedule. (5) Changes in state "double-dipping" laws – these laws prevent workers from receiving state pensions they have earned from SS-ineligible government work if they are taking any Social Security, thus effectively forcing many workers not to take Social Security benefits. (6) The "Special Minimum PIA" – this creates variation in effective marginal Social Security benefit rates for workers with a similar lifetime earnings but with different earnings histories. (7) Children's benefits – Children of retirees are eligible to receive 50% of the retiree's benefits, which creates variation in effective marginal Social Security benefits based on the parent's age at which a child was born.

individuals above certain income cutoffs ($44,000 for joint filers). If people recognize that benefits depend on contributions, this tax change should have led to labor supply responses that vary with the initial marginal benefit rate for those above the income cutoff but not for those below the income cut-off. To distinguish whether the response is due to an income or substitution effect, one can exploit an additional feature of the Social Security benefit rules: a couple has the choice of claiming either 150% of the benefits of the spouse with the highest benefits or of each spouse claiming 100% of their own benefits. Thus, this tax change will have only an income effect on individuals planning to claim spousal benefits, but an income and substitution effect for those claiming their own benefits. Moreover, the substitution effect will be stronger for those whose spouses will claim 50% of their benefits.

Third, there is variation due to rules on benefits for widows and widowers. An individual with a living spouse has the choice of claiming her own benefits or 50% of her spouse's benefits, while someone with a deceased spouse can choose between claiming her own benefits and 100% of her deceased spouse's benefits. Thus, individuals with a living spouse will claim their own benefits in a future year with the probability that the spouse is alive in that year and that one's own benefits exceed 50% of their spouse's benefits. However, upon the death of the spouse, only those with own benefits exceeding their spouse's benefits will continue to claim there own benefits. Thus, upon the death of a spouse, individuals with benefits that fall between 50% and 100% of their spouse's benefits will switch to claiming spousal benefits and they therefore no longer receive any marginal return from their Social Security contributions. Of course, the death of a spouse may also have a direct effect on labor supply. One can estimate the direct effect by looking at a second group of individuals whose marginal return from their Social Security contributions is not affected by the death of a spouse (i.e., those whose own benefits are less than 50%, or more than 100%, of spousal benefits). By comparing these two groups, it is possible to estimate the net effect of the marginal returns of Social Security contributions on labor supply.

These discontinuities should have a greater effect for those who had a higher marginal benefit rate beforehand. For instance, the reduction of the effective Social Security benefits due to the fact that the current year's earnings replace a prior year's earnings in the 35 highest years in the AIME formula has a greater impact on an individual who is on the 32% portion of the AIME-PIA schedule than on someone who is on the 15% portion. Hence, the identifying power of the regressions increases when we interact these sources of discontinuities with non-linearities in the Social Security benefit rules.

All these estimation techniques assume that individuals are rational forward-looking maximizers of expected life-time utility and that individuals recognize the stochastic nature of earnings and length of life. For example, we assume that individuals understand that with some probability they will receive future Social Security benefits based on their spouse's earnings (for example, if the spouse dies and has higher benefits than one one's own benefits).

3.2 Data

We perform our estimation using observations from the original cohort of the Health and Retirement Study (HRS), a longitudinal survey that can be linked to Social Security records. This cohort consists of individuals born between 1931 and 1941 as well as their spouses, who were born between 1900 to 1974 (with 90% between 1928 and 1947). Individuals are first interviewed in 1992 and have subsequently been reinterviewed every two years. Our data contains observations through 2004. In total, the original cohort of the HRS includes 13,367 individuals who are interviewed at least once.

Key to our analysis is the fact that we have historical Social Security earnings records for most individuals of the original cohort of the HRS. These records include yearly earnings (up to the Social Security contribution ceiling) from 1951 through 1991, as well as any Social Security benefits received. Additionally, the HRS contains self-reported earnings for years after 1991, which allows us to calculate expected Social Security wealth at each date that an individual is surveyed. Also, because the HRS contains information about both respondents and their spouses, we are able to exploit variation in the claiming of spousal benefits for our analysis.

The HRS also indicates the year and month that each individual retired (if they have done so when surveyed), allowing us to construct a precise measure of retirement status—both whether an individual is retired and for how long—at each survey date.

Furthermore, the HRS contains necessary control variables for our analysis, including total household assets, individuals' education, race, and Census region of residence.

In addition, the HRS is well suited for this study for two reasons beyond its wealth of data. The first is that the HRS focuses on older individuals. For these individuals, the majority of the earnings history is already in the past, and therefore they face much less uncertainty both about whether it will be optimal to claim spousal benefits upon retirement, as well as the level of earnings potentially displaced in the AIME formula by their current earnings. Moreover, because these individuals are near retirement, the present value of Social Security benefits is relatively high.

Second, existing empirical evidence suggests that retirement decisions may be more responsive to incentives than are the hours-of-work decisions of prime-aged workers. The HRS sample therefore involves exactly the age group that is most likely to respond to the marginal incentives provided by Social Security.

Table I shows summary statistics of our data. We include person-year observations of only those individuals who have no yet retired or who retired in the current year. We have a total of 4477 individuals in our sample, of whom 2337 are men and 2140 are women, and 26,195 person-year observations overall. Each year, an individual has approximately an 11% chance of retiring, a hazard rate that does not significantly vary by sex. Sixty-two percent of all individuals, 84% of men, and 37% of women report at least 35 years of earnings. The average age for individuals of both sexes is 60, and race and education do not vary significantly by gender. On average, individuals report that their earnings reached the Social Security contribution maximum in 7 years and that their spouse's did so in 3 years. However, for men this breakdown is 11 years and less than ½ year, respectively, and for women it is 1 year and roughly 6.5 years.

3.3 Social Security Wealth and the Effective Marginal Benefit Rate

Individuals' labor supply is expected to respond to the level of Social Security benefits (via the income effect) and to effective marginal Social Security benefits (via the substitution effect). We focus on estimating the effects of the effective marginal benefits on labor supply for two reasons: first, only changes in labor supply due to the substitution effect matter for efficiency, and, second, the Social Security benefits schedule primarily exhibits discontinuities in the marginal benefits, thus allowing for a more credible empirical identification of these effects. Since the effective Marginal Social Security Benefit Rate (*MBR*) is defined as the marginal effect of current labor supply on expected Social Security Wealth, we start by defining expected Social Security Wealth in year *t*, $E_t[SSW]$. We define Social Security Wealth for an individual and his or her spouse combined: conceptually, it is the expected present discounted value of all future net-of-tax payments from the Social Security Administration to the individual and his or her spouse.

Let $T(i)$ denote the year in which individual *i* reaches the full-benefit retirement age, let $\mathbf{y}_{it} = \{ \ldots, y_{it-2}, y_{it-1}, y_{it}\}$ denote the individual's earnings history up to and including year *t*, and let $\mathbf{y}_{it}^{+} = \{y_{it+1}, y_{it+2}, \ldots \}$ denote earnings in future years. Unless otherwise noted, all monetary variables are expressed in real 2003 dollars. Let $PIA_{iT}(\mathbf{y}_{it}, \mathbf{y}_{it}^{+})$ be the Primary Insurance Amount of individual *i* as a function of the individual's earnings profile. The PIA is determined by first calculating the AIME by taking the average of the 35 highest years of monthly indexed earnings and converting this AIME to a PIA using the AIME-PIA conversion schedule.

In year *t*, future earnings are not yet known. We therefore treat them as stochastic variables. In particular, we calculate the age- and gender-specific probability of future labor force participation based on the age- and gender-specific retirement hazard rates. We calculate expected future earnings conditional on being in the labor force by applying the age- and gender specific earnings growth to each year's earnings. Let r_{it} denote the probability that individual *i*'s earnings ended in year *t* (because of retirement or death).

An individual's Social Security benefits are adjusted based on his or her retirement age. Let $A(t,T)$ be the adjustment factor in year *t* for someone whose full-benefit retirement age is year *T*. This factor is zero in years before the early retirement age is reached. Finally, let variables that apply to the spouse be denoted with a superscript *s*. For individuals without a spouse, these take on a value of zero.

An individual whose spouse is alive has a choice of receiving benefits based on his or her own PIA or based on 50% of his or her spouse's PIA. Moreover, if the spouse is deceased, an invididual can claim benefits based on the maximum of the individual's own and spouse's PIA. Let p_{ti} denote the probability that individual *i* is alive in year *t*, based on the age- and gender-specific cohort life tables used by the Social Security Administration. Then the expected own Social Security benefit in year $t+k$ is found by taking the expectation over all possible combinations of the own and spousal retirement date of the maximum benefits in that year:

$$
\text{(1)} \quad \begin{aligned} E[SSB_{i,t+k} \mid \mathbf{y}_t, \mathbf{y}^s_t] = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} & r_{i,t+l} r^s_{i,t+m} A(t+k, T(i)) 1(l \le k) p_{i,t+k} \\ & \Big(p^s_{i,t+k} \ \text{Max}[EP_{it}(t+l), EP^S_{it}(t+\text{Min}[k,m])/2] + \\ & (1-p^s_{i,t+k}) \text{Max}[EP_{it}(t+l), EP^S_{it}(t+m)] \Big) \end{aligned}
$$

where $1(l<k)$ is the indicator function (which equals one if the expression between parentheses is true) and $EP_{it}(t+l) = \text{E}_t[PIA_{iT}(\mathbf{y}_{\text{it}}, \mathbf{y}_{\text{it+}})|r_{i,t+l}=1]$, that is, it is the expectation in year t of the PIA of individual i conditional on individual i retiring or dying in year $t+l$. Similarly, expected spousal Social Security benefits in year $t+k$ are given by:

$$
\text{(2)} \quad \begin{aligned} E[SSB^S_{i,t+k} \mid \mathbf{y}^s_t, \mathbf{y}_t] = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} & r^s_{i,t+l} r_{i,t+m} A(t+k, T(i)) 1(l \le k) p^s_{i,t+k} \\ & \Big(p_{i,t+k} \ \text{Max}[EP^S_{it}(t+l), EP_{it}(t+\text{Min}[k,m])/2] + \\ & (1-p_{i,t+k}) \text{Max}[EP^S_{it}(t+l), EP_{it}(t+m)] \Big) \end{aligned}
$$

Taking the present discounted value of the Social Security benefits of both spouses in future years yields expected Social Security Wealth (SSW):

$$
\text{(3)} \quad SSW_{it}(\mathbf{y}_t, \mathbf{y}^s_t) = \sum_{k=1}^{\infty} \Big(E[SSB_{i,t+k} \mid \mathbf{y}_t, \mathbf{y}^s_t] + E[SSB^S_{i,t+k} \mid \mathbf{y}^s_t, \mathbf{y}_t] \Big) (1+\rho)^{-k}
$$

where ρ is the real discount rate.

The Social Security schedule generally has different effects on the extensive and intensive margins of labor supply. We thus define two measures of the effective marginal Social Security benefit rate. To estimate the effects on the intensive margin, we define the *MBRI* as the change in Social Security wealth due to a marginal increase in the current year's earnings:

$$(4) \qquad MBRI_{it}(\mathbf{y}_t, \mathbf{y}_t^s) = \partial SSW_{i,t}(\mathbf{y}_t, \mathbf{y}_t^s) / \partial y_t$$

Because the *MBRI* is endogenous to the currrent's year earnings, we evaluate *MBRI* at a level of current earnings that is equal to last year's earnings.[3] To estimate the effect of the Social Security schedule on the incentive to retire, we calculate the *average* effective Social Security benefit rate if the individual retires at the very end rather than at the very beginning of the current year. As before, we evaluate this expression assuming that this year's earnings are equal to last year's earnings:

$$(5) \qquad MBRE_{it}(\mathbf{y}_t, \mathbf{y}_t^s) = \left(SSW_{it}(\mathbf{y}_t, \mathbf{y}_t^s \mid \mathbf{y}_{t+1}^+ = 0) - SSW_{it}(\mathbf{y}_t, \mathbf{y}_t^s \mid \mathbf{y}_t^+ = 0)\right) / \, y_{it}$$

3.4 Empirical Strategy

We employ two empirical strategies. In the first strategy, which we call the "Combined Regression Discontinuity Approach," we exploit all non-linearities and discontinuities in the Social Security schedule at the same time. We do this by running a regression of the following form:

$$(6) \qquad y_{it} = MBR_{it}(\boldsymbol{X}_{i,t-1})\,\alpha + SSW_{it}(\boldsymbol{X}_{i,t-1})\,\beta + f(\boldsymbol{X}_{i,t-1}, \boldsymbol{\gamma}) + \boldsymbol{Z}_{i,t-1}\,\boldsymbol{\delta} + \varepsilon_{it}$$

where y_{it} is a measure of labor supply (hours worked, labor earnings, retirement status) of individual i in year t, $\boldsymbol{X}_{i,t-1}$ is a vector of individual characteristics (including past earnings) that determine the marginal benefit rate, $\boldsymbol{Z}_{i,t-1}$ is a vector of other explanatory variables for labor supply while α, β, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are parameters to be estimated, and ε is an

[3] We do this in order to be safe. However, we suspect that, in practice, this endogeneity is less of a concern in this setting because variation in the current year's earnings is unlikely to move someone over one of the bend-points of the AIME-PIA schedule.

error term. When the outcome variable is retirement, the marginal benefit rate *MBR*(.) is equal to *MBRE* as defined by equation (5), while it equals *MBRI* as defined by equation (4) when the dependent variable is one of the other measures of labor supply. The functional form of the marginal benefit rate is given by the Social Security benefit formula and, critically, contains discontinuities. By contrast, the function *f*(.) is a *continuous* but flexible function of the same characteristics that determine the marginal benefit rate.[4] Thus, the parameter α, or the labor supply response to the the marginal Social Security benefit rate is identified by the jumps in the marginal benefit schedule. If this response is weaker than the estimate of the labor supply response to changes in income taxes as estimated by other studies, we would conclude that individuals only partially take the Social Security tax-benefit linkage into account. Since the level of Social Security wealth, *SSW*, contains very few discontinuities, the parameter β will likely not be estimated with much precision and *SSW* serves primarily as a control variable.

The obvious benefit of this strategy is that it is likely to have greater statistical power because it uses variation from all discontinuities and non-linearities in the Social Security benefit schedule. However, it also has some drawbacks. Because there are no objective standards as to what constitutes a sufficiently flexible form for the non-linear continuous control function *f*(.), some of the identification of the parameter of interest may come from the functional form rather than from the discontinuities (to the extent that we know the correct functional form of the Social Security benefit schedule, this is less of a drawback). Furthermore, by combining all discontinuities in the same framework, we do not know which of the discontinuities drive the results.

In order to better understand which of the discontinities drive the results from the Combined Regression Discontinuity Approach, we also employ a second strategy, which we call the "Single Discontinuity Approach." In this strategy, we focus on one discontinuity at a time, and use variation *only* from that discontinuity to identify the effect of the marginal benefit rate on labor supply. We will apply this approach to each of the three main discontinuities: variation in earnings histories, the 1993 tax reform, and death of a spouse.

---

[4] Coile and Gruber (2000) also use this strategy.

The benefit to the Single Discontinuity Approach is that we know the source of variation driving the estimate. As a result, we can be confident that the labor response estimate is not identified off of functional form. The drawback of this approach is that it has less statistical power because it ignores other sources of variation in the Social Security benefit schedule.

3.5 Single Discontinuity Approach Specifications

Specification A: Exploiting Variation in Timing of Earnings

If, as is the case for the vast majority of older workers, current earnings exceed the 35^{th} highest year of past earnings, then the fact that current earnings may or may not replace a prior year of earnings in the AIME calculation only affects the extensive margin of labor supply. In other words, it can affect an individual's decision of whether or not to retire, but conditional on not retiring, it will not affect this individual's incentive to work or earn more or less. We will therefore focus on the discontinous effect of earnings history on Social Security benefits to estimate the effects of the Social Security tax-benefit link on retirement decisions. We will assume that retirement is an absorbing state.

We model each individual's decision problem as one in which the individual decides at the start of each year whether or not to retire. The individual would need to assess to what extent working an extra year will increase future Social Security benefits. For individuals whose earnings in the 35^{th} highest year of earnings are zero, an additional year of earnings will increase the person's AIME by next year's indexed earnings divided by 35: , $\Delta AIME = \tilde{y}_{i,t+1} / 35$, with the tilde denoting that earnings are indexed. In contrast, someone whose earnings in the 35^{th} highest year of earnings exceed zero, will experience a smaller increase in his or her AIME because next year's earnings replace an earlier year of earnings. Thus for this person, an additional year of earnings will increase his or her AIME by $\Delta AIME = (\tilde{y}_{i,t+1} - \tilde{y}_i^{35}) / 35$, with $\tilde{y}_i^{35}$ denoting the 35^{th} highest year of indexed earnings. Thus, the fact that the Social Security benefits only depend on the 35 highest years of indexed earnings creates a difference in retirement incentives between these two individuals. This difference is given by $\tilde{y}_i^{35} / 35$ in absolute terms, or by $\tilde{y}_i^{35} / (35\,AIME)$ in relative terms. We create four variables that proxy for these retirement incentives:

a. I($\tilde{y}_i^{35} > 0$), which is a dummy variable that equals one if the 35th highest year of indexed earnings is greater than zero.

b. I($\tilde{y}_i^{35} > 0.2\,\bar{\tilde{y}}_{t \text{ of } y^{35}}$), which is a dummy variable that equals one if the 35th highest year of indexed earnings exceeds 20% of the mean indexed earnings in the year when those earnings took place. By setting the threshold somewhat above zero, we ensure that our estimates are not affected by those whose 35th highest year of earnings is approximately zero (but not quite zero, e.g., due to a summer job in high school or college).

c. I($\tilde{y}_i^{35} > 0.5\,\bar{\tilde{y}}_{t \text{ of } y^{35}}$). This proxy is similar to the second one, except that the earnings threshold is higher.

d. $\tilde{y}_i^{35} / \sum_{k=1}^{5} (\tilde{y}_{i,t-k} / 5)$, which expresses the displaced indexed earnings as a fraction of average indexed earnings in the past 5 years (which, in turn, will be highly correlated with the person's AIME).

Let the retirement incentive, as measured by one of the these four proxies, be denoted by *I*. To estimate the effect of this incentive on an individual's retirement decision, we first run a pooled OLS linear probability model of the form:

$$(7) \qquad R_{it} = I_{it}\,\alpha + \boldsymbol{Z}_{i,t-1}\,\boldsymbol{\gamma} + \varepsilon_{it}$$

where where R_{it} is a dummy variable whose value is one if individual *i* retired in year *t*. The vector $\boldsymbol{Z}$ contains an extensive set of individual and spousal characteristics that also influence the retirement decision. By including an broad set of controls, we ensure that the coefficient on the retirement incentive variable is unlikely to be driven by other characteristics of the individual that might also influence the retirement decision. These controls include a full set of own age by gender dummies, a set of dummies capturing all possible combinations of own age category, as well as own gender and spousal age category (age categories: < 62, 62-65, >65). They also include a cubic polynomial in the log of the present discounted value of lifetime social security earnings, the log of the present discounted value of spousal lifetime social security earnings, the variance of the top 20 years of earnings, the number of years in the Social Security benefits formula in which the person's earnings were equal to the Social Security earnings maximum, and an analogous variable for the spouse.[5] Additionally, we control for the log of household assets, own and spousal education dummies (high school dropout, high school graduate, some college, bachelor's degree or more), own race-ethnicity dummies (non-Hispanic black, non-Hispanic white, Hispanic, other), and dummies for the 10 Census regions. In addition, we include the log of the present discounted value of the last 5 years of earnings when the the proxy for the retirement incentive is a dummy variable. When the proxy for the retirement incentive is the ratio variable, we include a control for the inverse of the log of the present discounted value of the last 5 years of earnings. This ensures that the coefficient on retirement incentive ratio variable is driven by variation in $\tilde{y}_i^{35}$ rather than by variation in the present discounted value of the last 5 years of earnings. We cluster all standard errors by individual.

We estimate this regression on a sample of individuals who are either currently working or who retired in the current year. We exclude the following individuals: (i) those with earnings histories with fewer than 20 years of positive indexed earnings, as they likely differ on many unobservable dimensions from those who are replacing positive earnings, (ii) those who are younger than 55, because they may be retiring for very idiosyncratic reasons, and (iii) those claiming not to be retired but with combined earnings less than \$5000 (in 2003 real dollars) in the past 5 years (because either their retirement variable is mismeasured or because they are working in a sector not covered by Social Security).

We first estimate pooled OLS linear probability models because of their simplicity and transparency. However, given that retirement is an absorbing state for most individuals, we also estimate the the effect of each of our proxies for the retirement incentive in a Cox proportional hazard model.

[5] We dummy out all the logs of dollar amounts if the dollar amount is less than \$1000 in 2003 real dollars.

Specification B: Exploiting Variation from the Widow/Widower Benefit Rules

Individuals have the choice of claiming their own benefits or 50% of their spouse's benefits when the spouse is alive, but can choose between their own benefits and 100% of the spouse's benefits when the spouse is deceased. These rules imply that the death of a spouse makes it optimal for some individuals to claim spousal benefits rather than own benefits. As a result, these individuals experience a sudden drop in the marginal Social Security benefit rate. Similarly, an individual whose spouse dies will experience a drop in his marginal benefits if the spouse was planning to claim 50% of his benefits, but no change in marginal benefits if the spouse was planning to claim her own benefits. The death of a spouse will also affect Social Security wealth directly, but, as the examples above illustrate, the direct effect on Social Security wealth is often different from the effect on marginal returns. Finally, the death of a spouse may of course also directly affect labor supply.

For all individuals married in year *t*, we predict how the death of one's spouse in year *t*+1 would affect Social Security Wealth and the Marginal Benefit Rate. These predicted changes are based only on information available in year *t*. Specifically, to calculate the predicted change in Social Security Wealth, $\Delta SSW_{it}(\mathbf{y}_t, \mathbf{y}_t^s)$, we first calculate SSW as defined by equation (3) using $p_{i,t+k}^s$, the spousal probability of living in year *t+k*, from the cohort life tables and using $r_{i,t+k}^s$, the spousal probability of retiring in year *t+k*, from the empirical retirement hazard rate distribution. Next, we use equation (3) to calculate SSW conditional on the spouse dying in year *t*+1 by setting $p_{i,t+k}^s$ equal to zero for all *k*>1 and by setting $r_{i,t+1}^s$=1 if the spouse was not yet retired. The difference between the unconditional SSW and the SSW conditional on the spouse dying in the next year yields $\Delta SSW_{it}(\mathbf{y}_t, \mathbf{y}_t^s)$. We then calculate the change in the marginal benefit on the intensive margin by:

(8) $$\Delta MBRI_{it}(\mathbf{y}_t, \mathbf{y}_t^s) = \partial \Delta SSW_{i,t}(\mathbf{y}_t, \mathbf{y}_t^s) / \partial y_t,$$

and the change in the marginal benefit rate on the extensive margin by:

(9) $$\Delta MBRE_{it}(\mathbf{y}_t, \mathbf{y}_t^s) = \left(\Delta SSW_{it}(\mathbf{y}_t, \mathbf{y}_t^s \mid \mathbf{y}_{t+1}^+ = 0) - \Delta SSW_{it}(\mathbf{y}_t, \mathbf{y}_t^s \mid \mathbf{y}_t^+ = 0)\right) / \, y_{it}$$

We will use the symbol ΔMBR_{it} to denote $\Delta MBRI_{it}$ when we use hours or earnings as our measure of labor supply and $\Delta MBRI_{it}$ when we use labor force participation as our measure of labor supply.

Let $\Delta y_{i,t+2} \equiv y_{i,t+2} - y_{i,t}$ denote the change in labor supply between year $t+2$ and year t, and let $SDied_{t+1}$ be a dummy variable that equals one if the spouse died in year $t+1$. The following OLS regression relates an individual's change in labor supply in the years surrounding the possible death of a spouse to the incentive and income effects from the Social Security system:

(10) $$\Delta y_{i,t+2} = SDied_{t+1} \times \Delta MBR_{it}\, \beta_1 + SDied_{t+1} \times \Delta SSW_{it}\, \beta_2 + SDied_{t+1}\, \beta_3 + \Delta MBR_{it}\, \beta_4 + \Delta SSW_{it}\, \beta_5 + \boldsymbol{X}_{it}\, \gamma + \delta_t + \varepsilon_{it}\,,$$

where $\boldsymbol{X}_{it}$ is a vector of demographic controls (including a full set of gender×age dummies), δ_t is a full set of year dummies, and ε_{it} is the error term. This equation is estimated on a sample of individuals whose spouse either died in year $t+1$ or did not die between the years of $t-1$ and $t+3$. Including individuals whose spouse did not die helps to alleviate concerns that changes in labor supply might be spuriously related to ΔMBR_{it} because ΔMBR_{it} is a non-linear function of own demographic characteristics.

The labor supply response to the change in the Marginal Benefit Rate is estimated by the coefficient β_1, while β_2 estimates the response to the change in Social Security Wealth. The direct impact of a spousal death on labor supply, controlling for any wealth or incentive effects from Social Security, is given by β_3. The remaining coefficients do not have an economic interpretation because they are coefficients on control variables.

4. Results

4.1 The Effect of the Maximum-35-Years Rule on Retirement

Since the key outcome variable is retirement, we show the retirement hazard rate by gender in Figure I. The figure shows that there is a considerable age range within which retirement hazard rates are substantial. We find that for both men and women the retirement hazard rate increases sharply at age 62 and at 65, which are the early and full-benefit Social Security retirement ages.[6,7] In addition, there appears to be a noticeable increase in the retirement hazard at age 60 for women and at age 69 for men.

Recall that our key explanatory variable measures the extent to which current earnings replace a previous year's earnings in the 35 highest years of indexed earnings in the Social Security benefit formula. The most basic proxy for this variable is a dummy variable that equals one if the earnings being replaced are strictly positive. Figure II graphs the hazard rates by age and gender for this dummy variable becoming 1. In other words, it shows for each age and gender cell, what fraction of people who previously did not have an incentive to retire because of the maximum-35-year rule, now, for the first time, face this incentive. The figure shows that this hazard rate is high for a large age range. This is reassuring because it means that the variation due to this rule is unlikely to be highly correlated with other age-specific incentives to retire. Figure III is similar to Figure II, except that it measures the retirement incentive from the maximum-35-year rule by a dummy that equals 1 if the earnings being replaced by the current earnings are greater than 20% of the mean earnings in the year they were earned. This figure confirms the key message from Figure II: there is ample variation in the age at which the retirement incentives start to increase due to the maximum-35-year rule.

Figure IV plots the retirement hazard rates by the number of years that the individual has faced increased retirement incentives due to the maximum-35-year rule. For example, the retirement hazard rate in year 1 is the retirement hazard rate in the first year in which current earnings replace previous earnings in the Social Security benefit formula. Similarly, the hazard rate in year -5 denotes the hazard rate in the year when the

---

[6] The full-benefit retirement age is gradually rising to 67. However, when the HRS was conducted the full-benefit retirement age was 65 for most individuals in our sample.

[7] We do not show retirement hazards beyond the age of 69 because we have fewer than 100 observations in each age-gender cell off of which to estimate hazard rates for ages 70 or greater.

individual would still have to work 5 more years before current earnings would start to replace previous earnings in the Social Security benefit formula.

Figure IV shows that the retirement hazard rate for men starts to increase sharply in year 1: from years -10 to year 0, the retirement hazard rate is roughly constant at about 1%, while starting at year 1, the retirement hazard increases at roughly 2% *per year*. A basic economic model of retirement would predict that the *level of* the retirement hazard would increase when retirement incentives rise, not that the *rate at which* the retirement hazard increases becomes larger when retirement incentives are stronger. Thus, in the framework of a basic economic model of retirement, this figure is not direct evidence that individuals respond to retirement incentives from the Social Security rules because we do not observe a large jump in the level of the retirement hazard around year 0. However, it seems plausible that retirement incentives could affect the rate at which the retirement hazard increases in a more sophisticated model of retirement.[8] In any case, the figure raises the question of why the retirement hazard rate starts to increase exactly in the year when the maximum-35-year rule becomes binding. Though it is possible that this is a coincidence, the maximum-35-year rule remains a prime suspect.

For women, we see a similar, although somewhat less pronounced, pattern. Their retirement hazard increases at a relatively low rate (about 0.3% per year) in the years before the maximum-35-year rule binds, and starts to increase at a much higher rate (about 1.5% per year) around the year that the maximum-35-year rule becomes binding. Figure IV shows that we obtain very similar results if we treat prior years' earnings as zero if they were less than 20% of mean earnings in the years they were earned. In other words, the results are not sensitive to prior years with small but positive earnings.

Age and the number of years for which the maximum-35-year rule is binding are positively correlated. One might therefore be concerned that the relationship between the number of years that the maximum-35-year rule is binding and the retirement hazard rate is driven by the fact that those who have worked longer are more likely to retire simply because of age. To address this concern, we regress a dummy for retiring on a full set of

[8] For example, as individuals continue to work beyond 35 years, they crowd out years with progressively higher earnings. The incentive to retire therefore generally rises with the number of years worked beyond 35. Alternatively, a model in which it takes individuals time to learn about or react to retirement incentives could also potentially explain the pattern in this graph.

age dummies and on a set of dummies for the number of years that the maximum-35-year rule has been binding. We run these regressions on a sample of individuals who are working or retire in the current year, and run them seperately for men and women. Figure VI shows the resulting coefficients on the dummies for the number of years that the maximum-35-year rule has been binding. The figure shows that, after controlling for age, the retirement hazard rate for men is roughly constant before year 1, the year when the maximum-35-years rules starts to bind. After year 1, however, the retirement hazard rate starts to increase sharply. Thus, after controlling for age, we obtain the same qualitative finding as in figure IV: retirement hazard rates start to increase sharply as soon as current earnings start to replace previous earnings in the Social Security benefit formula. For women, the findings are less pronounced, but their retirement hazard rate also appears to be higher in the years when the maximum-35-years rule binds.

While these graphical results are highly suggestive, they do not control for many additional individual characteristics that might both affect the retirement decision and be correlated with whether the maximum-35-year rule is binding. In order to control for these characteristics, we run regressions as defined by equation (7) in section 3.5. Table II shows results from this regression for men and women combined and by gender. The outcome variable is a dummy for retiring, while the key explanatory variable is a dummy that that equals one when the maximum-35-year rule is binding (i.e., that equals one if the 35^{th} highest year of prior earnings is positive). We find no significant effect of this dummy variable on the retirement hazard rate for any of the three samples. The standard errors are reasonably small, around 0.006 for the combined sample and 0.009 for each gender separately. Given that the mean retirement hazard rate in our sample is 0.11, this seems to imply that we can rule out a large effect of the maximum-35-year rule on retirement decisions. The log of the PDV of total Social Security earnings, as well as the cube thereof, have a negative effect on retirement, while the square has a positive effect. The log of total spousal Social Security earnings has a positive effect for men and a negative effect for women. Individuals with more assets are more likely to retire. Furthermore, the more education an individual has, the less likely they are to retire.

The first row of Table III shows the coefficient on the key explanatory variable of the three regressions shown in Table II. In these regressions, we measure the effect of the maximum-35-year rule by a dummy that equals one if the current year's earnings replace any non-zero amount of prior earnings in the Social Security benefit formula. A potential problem with this proxy is that the retirement incentive is rather small if the current year's earnings replaces prior earnings that were non-zero yet small. We thus refine our proxy by requiring that the prior year's earnings that are being replaced exceed a certain threshold. We choose two earnings thresholds that we believe rule out trivial amounts of prior year's earnings: at 20% and 50% of the mean earnings in the year the earnings took place. Rows (2) and (3) of table III show the results of the effect of the maximum-35-year rule on the retirement probability if we impose the 20% and the 50% thresholds. We now find that individuals are 1.3 percentage points more likely to retire if the current year's earnings replace a previous year's earnings that meets the 20% threshold. Given that the retirement hazard rate is 0.11 on average, a 1.3 percentage point increase corresponds to a 12 percent increase. This effect is significant at the 5% level for the combined sample of men and women. It is not statistically significant for men and women separately, though the point estimates are similar in magnitude. The effect on retirement of the maximum-35-year rule rises to 3.1 percentage points when we apply the 50% threshold on prior earnings. This effect is significant at the 1% level and is driven by both men and women. In row (4), we measure the impact of the maximum-35-year rule by the earnings in the 35th highest prior year as a fraction of average earnings in the last 5 years. We suspected that this would be the most accurate measure of the effect of the maximum-35-year rule because it scales rather than dichotomizes the earnings in the 35th highest year. We, however, fail to find any significant effect of the maximum-35-year rule using this variable. The point estimates are actually negative, though nowhere near statistical significance. In future research, we plan to explore the reasons that our results are so sensitive to the exact variable chosen to proxy for the impact of the maximum-35-year rule in the Social Security benefit formula.

Rows (5) through (12) of Table III split the results by education. We label those with only a high school degree or less as "low education". Those with low education comprise 58% of the men in our sample, 60% of the women, and 59% of the overall sample. We split the sample by education because education is a consistent predictor of lifetime earnings. Those with lower levels of lifetime earnings are more likely to be on one of the steeper segments of the AIME-PIA schedule. As a result, their Social Security benefits tend to be more sensitive to the maximum-35-year rule. Moreover, for those with lower earnings, Social Security tends to comprise a larger fraction of retirement income, thereby plausibly increasing the effect of the maximum-35-year rule on the retirement decision. We would therefore expect our results to be more pronounced for the low education group. Table III, however, shows that the retirement response appears quite similar for men of each education group. While this could be explained by the low education group being less responsive to an equally sized incentive (e.g. because they have a less clear understanding of the Social Security rules), the lack of a differential effect by education does raise some doubts about whether the significant coefficients reflect a causal effect of Social Security rules on retirement decisions.

However, there does appear to be a larger response by women of lower education than by women of high education, in line with our prediction. Still, standard errors are sufficiently large for educated women that while we can rule out a response of the magnitude of their less educated counterparts, we cannot rule out a response that is as large as the one educated men display.

Table IV shows the same specifications as in Table III, except that all models are specified as Cox hazard models rather than as pooled OLS linear probability models. The results in table IV are generally consistent with those in table III but the standard errors tend to be larger. As a result, fewer coefficients are statistically significant. In fact, only when we apply the 50% earnings threshold, do we find a significant effect of the maximum-35-year rule on the retirement hazard for the whole sample. It is significant at the 1% level for all individuals, at the 5% level for men only, and at the 10% level for women only. In addition, consistent with our Table III results, we find that the 50% earnings threshold is significant at the 5% level for all low education individuals as well as for low education women, and at the 10% level for high education men, but not for high education women. Because the coefficients in a hazard model measure the impact of the explanatory variable on the log of the hazard ratio, they should be interpreted as the effect on the retirement hazard in percentage terms, not in percentage point terms (as was the case in the linear probability model).

5. Discussion

In this paper, we show how discontinuities in the Social Security benefit schedule can be exploited to estimate how individuals respond to the labor supply incentives provided by the Social Security benefit schedule. We identify three discontinuities that can potentially provide sufficient, credible variation in labor supply incentives: discontinuities due to the fact that only the 35 highest years of indexed earnings are included in the Social Security benefit formula, discontinuities due to the treatment of widow and widower benefits, and discontinuities due to a tax reform in 1993. We show how all sources of discontinuities can be combined with non-linearities in the Social Security benefit formula to yield a powerful estimator of the labor supply effects of the benefit rules. We also show how to exploit variation in marginal benefit rates due to each discontinuity separately in order to estimate the effect of the Social Security benefit formula on labor supply.

Our empirical results are based on variation in marginal benefit rates due to one discontinuity – the maximum-35-year rule. In future research, we will also use the variation from the other two main discontinuities. In addition, we will run specifications in which we exploit variation from all the discontinuities and non-linearities combined.

Graphs of the retirement hazard rate as a function of the number of years during which the maximum-35-year rule of the Social Security benefits formula binds show a remarkable pattern. The retirement hazard ratio starts to increase sharply in the year that the maximum-35-year rule becomes binding. This result is highly suggestive of the incentives implicit in Social Security benefit formula affecting individuals' retirement decisions. We also explore this relationship in a regression framework. The benefit of using regression analysis is that it allows us to include a rich set of controls that might otherwise confound the relationship between the maximum-35-year rule being binding and the decision to retire. The regression results show a more mixed picture. In some specifications, the retirement hazard increases significantly when the maximum-35-year rule becomes binding, that is, when the marginal effective Social Security benefit rate falls. In other specifications, however, we do not find a significant effect.

Taken together, we believe our findings are suggestive of people reacting to incentives in the Social Security benefit rules. However, until further research illumintes why only some specifications yield significant results and provides estimates based on other discontinuities, these findings remain tentative.

References

Coile, Courtney C. and Jonathan Gruber. 2000. "Social Security and Retirement." NBER Working Paper #7830.

Diamond, Peter and Jonathan Gruber. 1999. "Social Security and Retirement in the U.S.," in *Social Security and Retirement Around the World*, J. Gruber and D. Wise, eds. Chicago: University of Chicago Press.

Goda, Gopi Shah, John B. Shoven, and Sita Nataraj Slavov. 2006. "Removing the Disincentives in Social Security for Long Careers," Unpublished manuscript, Stanford University.

Feldstein, Martin S. and Jeffrey B. Liebman. 2002. "Social Security," in Alan J. Auerbach and Martin Feldstein, eds., *Handbook of Public Economics*, volume 4, Amsterdam: Elsevier.

Feldstein, Martin S. and Andrew A. Samwick. 1992. "Social Security Rules and Marginal Tax Rates," *National Tax Journal*, 45, pp. 1-22.

Gustman, Alan L. and Thomas L. Steinmeier. 2005. " The Social Security Early Retirement Age in a Structural Model of Retirement and Wealth", *Journal of Public Economics*, 89, pp. 441-463.

Kotlikoff, Laurance. 1996. "Privatizing Social Security: How It Works and Why It Matters," *Tax Policy and the Economy*, 10, pp. 1-32.

Krueger, Alan B. and Jörn-Steffen Pischke. 1992. "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation," *Journal of Labor Economics*, 10(4), pp. 412-437.

Laitner, John and Dan Silverman. 2006. "Social Security Reform: Changing Incentives for When to Retire," unpublished manuscript, University of Michigan.

Samwick, Andrew A. 1998. "New Evidence on Pensions, Social Security, and the Timing of Retirement," *Journal of Public Economics*, 70(2), pp. 207-236.

Stock, James H. and David A. Wise. 1990 "Pensions, the Option Value of Work, and Retirement," *Econometrica*, 58(5), pp. 1151-1180.

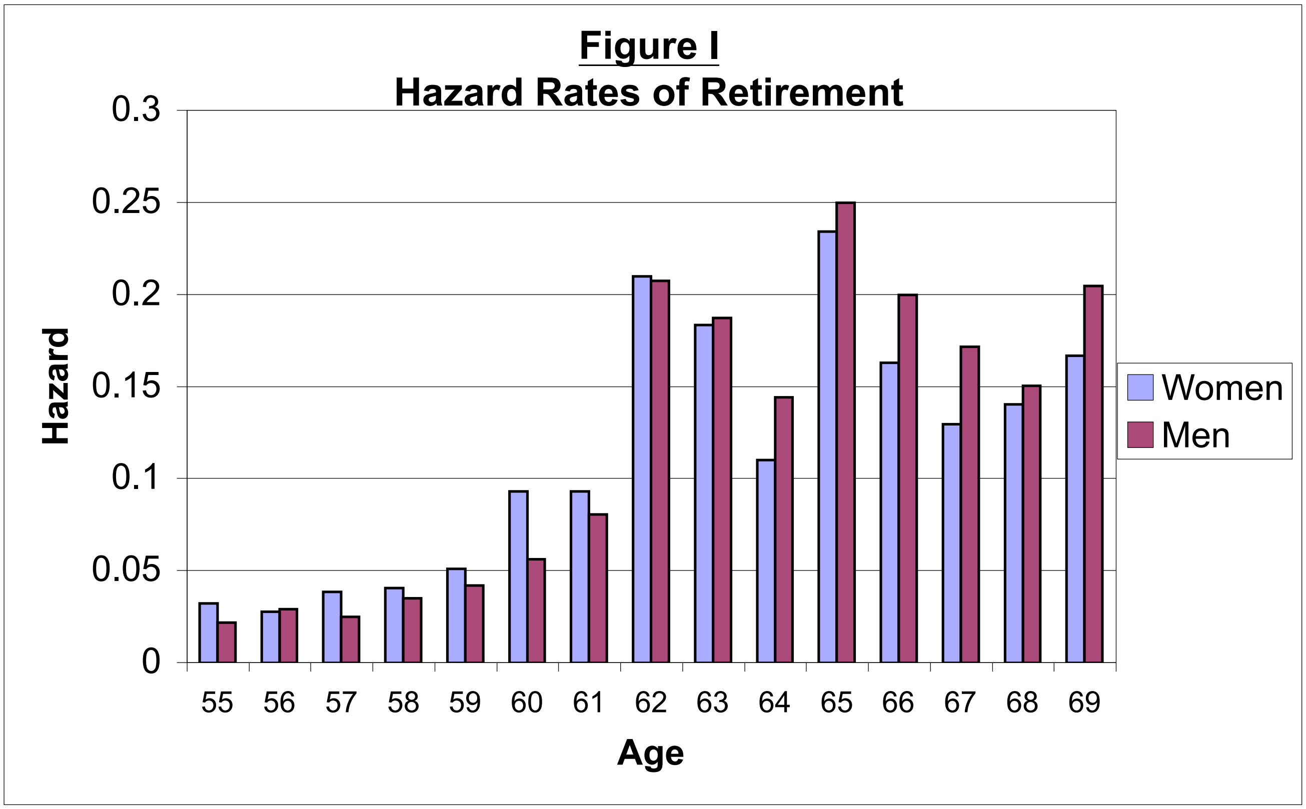
Figure I
Hazard Rates of Retirement
Hazard
0.3
0.25
0.2
0.15
0.1
0.05
0
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
Age
Women
Men

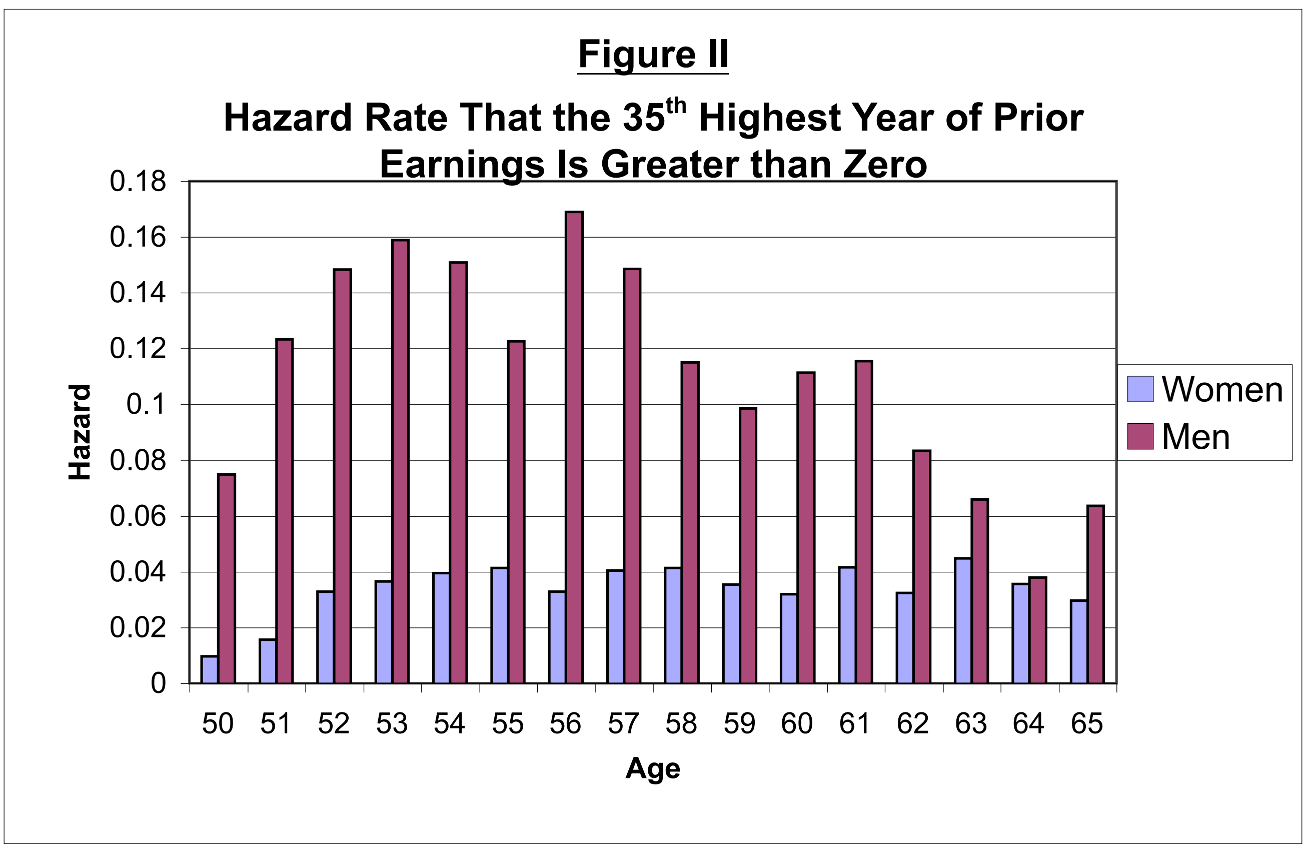
Figure II
Hazard Rate That the 35th Highest Year of Prior Earnings Is Greater than Zero
Hazard
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
Age
Women
Men

Figure III

Hazard Rate That the 35th Highest Year of Prior Earnings Is Greater than 20% of Average Wage

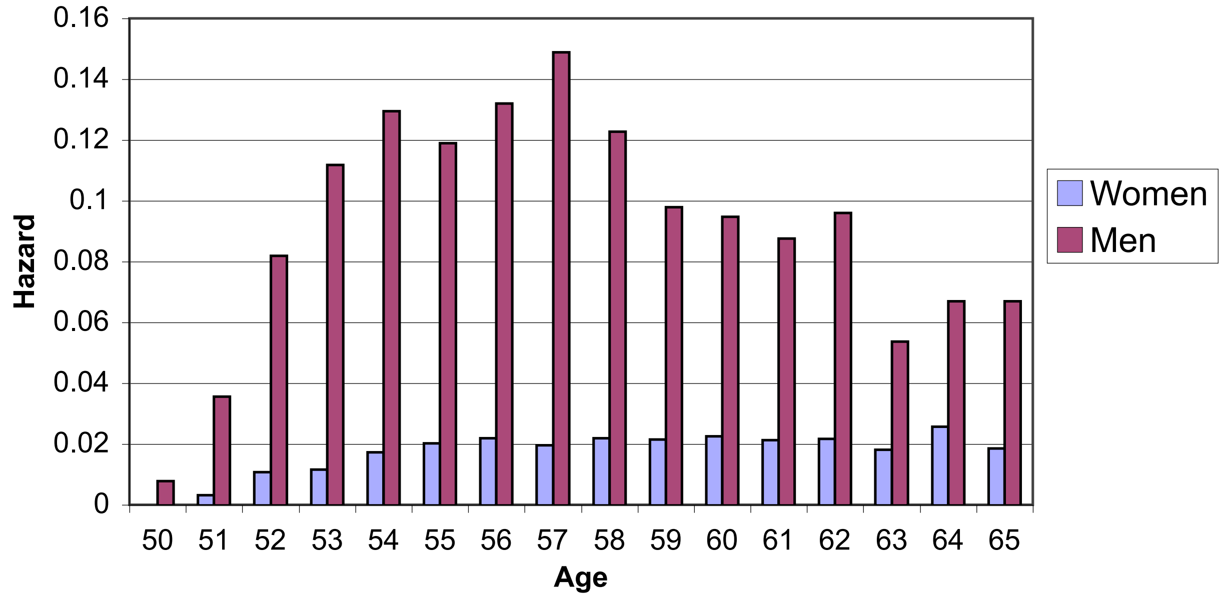

Figure IV

Hazard Rate of Retirement by Number of Years That the "Maximum-35-Years" Rule Has Been Binding

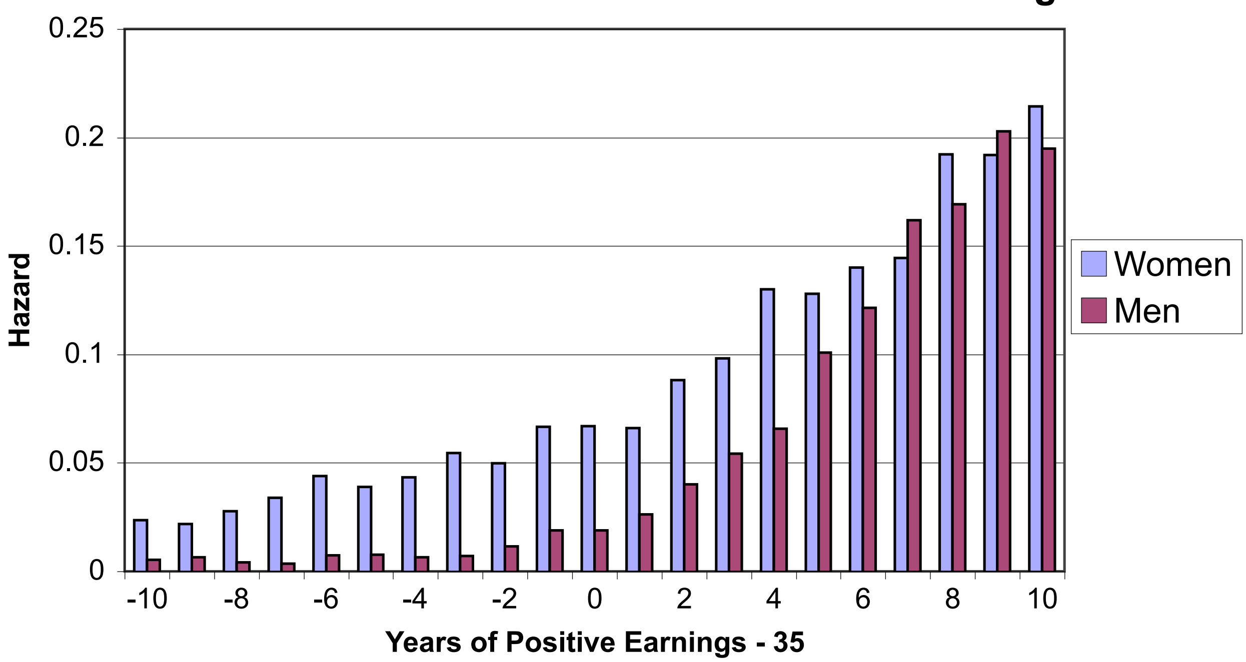

Figure V

Hazard Rate of Retirement by Number of Years That the "Maximum-35-Years" Rule Has Been Binding

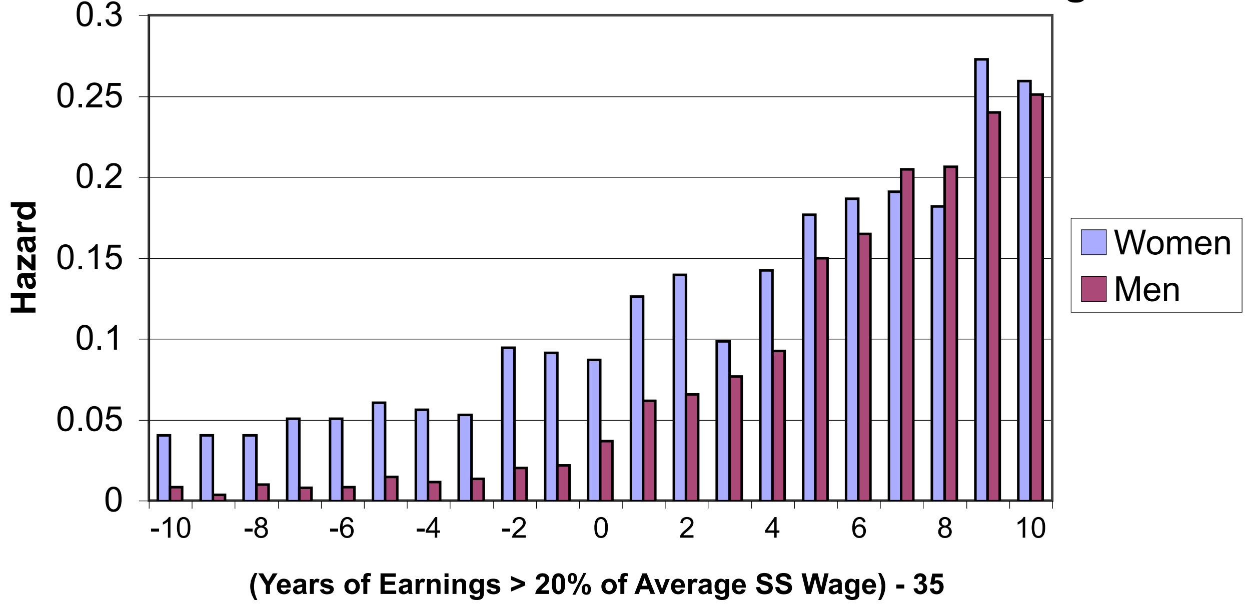

Figure VI

Hazard of Retirement by Earnings History After Controlling for Age

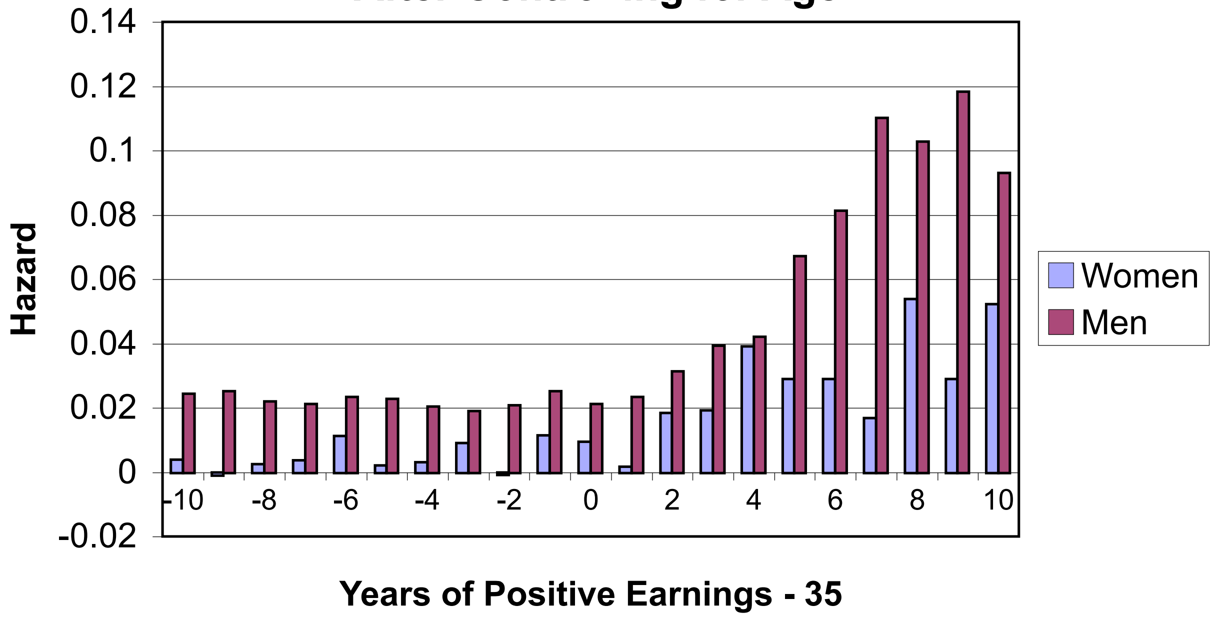

TABLE I--Summary Statistics

Variable	All Individuals Mean	All Individuals Std. Dev.	Men Mean	Men Std. Dev.	Women Mean	Women Std. Dev.
Retire This Year	0.110	(0.313)	0.113	(0.316)	0.107	(0.309)
Year	1997.1	(3.443)	1996.7	(3.335)	1997.5	(3.507)
35th Highest Year of Earnings	9221	(11137)	14821	(11680)	3001	(6065)
35th Highest Year of Earnings > 0	0.616	(0.486)	0.836	(0.371)	0.372	(0.483)
35th Highest Year of Earnings > 20%	0.449	(0.497)	0.694	(0.461)	0.176	(0.381)
35th Highest Year of Earnings > 50%	0.273	(0.445)	0.460	(0.498)	0.064	(0.246)
35th Highest Year / Avg. of Last 5 Years	0.232	(0.271)	0.354	(0.278)	0.095	(0.185)
Sex	0.526	(0.499)	1	(0)	0	(0)
Age	60.026	(3.459)	60.219	(3.513)	59.811	(3.386)
Variance of Top 20 Years' Earnings / 10^6	7.47	(8.34)	7.60	(7.41)	7.32	(9.26)
Log(Avg. PDV of Last 5 Years' Earnings)	10.310	(0.770)	10.495	(0.751)	10.103	(0.738)
Log(PDV of Total SS Earnings)	14.173	(0.664)	14.561	(0.452)	13.743	(0.592)
[Log(PDV of Total SS Earnings)]^2	201.327	(18.517)	212.236	(12.885)	189.210	(16.135)
[Log(PDV of Total SS Earnings)]^3	2866	(389)	3096	(276)	2610	(331)
Log(PDV of Spouse's Total SS Earnings)	9.596	(6.265)	10.743	(5.036)	8.321	(7.182)
Log(Household Assets)	11.117	(3.083)	11.341	(2.914)	10.869	(3.242)
Household Assets > $1000	0.942	(0.234)	0.951	(0.216)	0.932	(0.251)
Years of Maximum SS Earnings	6.771	(9.583)	11.995	(10.423)	0.969	(3.034)
Spouse's Years of Maximum SS Earnings	3.357	(7.316)	0.487	(1.806)	6.544	(9.489)
High School Dropout	0.199	(0.399)	0.216	(0.411)	0.181	(0.385)
High School Graduate (only)	0.390	(0.488)	0.360	(0.480)	0.423	(0.494)
Some College, No Bachelor's Degree	0.212	(0.408)	0.193	(0.394)	0.233	(0.422)
Bachelor's Degree or Higher	0.199	(0.399)	0.232	(0.422)	0.163	(0.370)
White	0.784	(0.411)	0.810	(0.392)	0.755	(0.430)
Non-Hispanic Black	0.136	(0.343)	0.106	(0.307)	0.169	(0.375)
Hispanic	0.064	(0.245)	0.068	(0.252)	0.059	(0.236)
Other Race	0.016	(0.126)	0.016	(0.126)	0.016	(0.127)
Years of Positive SS Earnings	35.540	(6.812)	38.657	(5.497)	32.079	(6.451)
Number of Observations	26195		13784		12411	
Number of Individuals	4477		2337		2140	

Notes: Observations are person-year
All log variables are set to zero for values < $1000 (in 2003 real dollars)
Sample is limited to individuals age 55+, with at least 20 years of labor force experience, with at least $5000 of earnings in the preceding 5 years, and who enter the year non-retired.
All spousal variables are set to zero for unmarried individuals.
Ratio of 35th Highest Year / Avg. of Last 5 Years is topcoded at 1.

Table II--OLS Change in Retirement Probability

Dependent variable: Dummy for retiring

Independent Variable	All Individuals		Men		Women	
35th Highest Year > 0	-0.001	(0.006)	-0.006	(0.009)	0.002	(0.008)
Variance of Top 20 Years' Earnings / 10^6	-0.005**	(0.003)	-0.010**	(0.004)	-0.003	(0.003)
Log(PDV of Total SS Earnings)	-5.537**	(2.804)	-0.025	(5.677)	-3.680	(4.375)
[Log(PDV of Total SS Earnings)]^2	0.388*	(0.205)	-0.001	(0.407)	0.244	(0.325)
[Log(PDV of Total SS Earnings)]^3	-0.009*	(0.005)	0.000	(0.010)	-0.005	(0.008)
Log(PDV of Spouse's Total SS Earnings)	-0.000	(0.003)	0.005*	(0.003)	-0.025***	(0.008)
Log(Household Assets)	0.008***	(0.002)	0.010***	(0.002)	0.008***	(0.002)
HH Assets > $1000	-0.083***	(0.020)	-0.090***	(0.029)	-0.079***	(0.028)
Years of Maximum SS Earnings/100	-0.035	(0.044)	0.015	(0.050)	-0.158	(0.131)
Spouse's Years of Maximum SS Income/100	0.099**	(0.041)	-0.081	(0.196)	0.204***	(0.052)
High School Graduate (only)	-0.015***	(0.006)	-0.010	(0.008)	-0.023***	(0.008)
Some College, No Bachelor's Degree	-0.033***	(0.007)	-0.027***	(0.009)	-0.041***	(0.009)
Bachelor's Degree or Higher	-0.042***	(0.008)	-0.046***	(0.010)	-0.042***	(0.011)
Non-Hispanic Black	0.007	(0.006)	-0.003	(0.010)	0.010	(0.008)
Hispanic	-0.014*	(0.008)	-0.019*	(0.012)	-0.017	(0.011)
Other Race	0.007	(0.011)	0.021	(0.016)	-0.005	(0.015)
Full set of gender-specific age dummies	Yes		Yes		Yes	
Full set of dummies for gender×(own age category)×(spousal age category)	Yes		Yes		Yes	
N (person-year observations)	26195		13784		12411	
Clusters (Number of People)	4477		2337		2140	
R^2	0.0613		0.0662		0.0608	

Notes: Standard errors are in parentheses and clustered by individual.

Sample restricted to non-retired individuals aged 55+ with at least 20 years' work experience and at least $5000 earned income in preceding 5 years.

* indicates p-value < .10

** indicates p-value < .05

*** indicates p-value < .01

Table III--OLS Change in Retirement Hazard

Dependent variable: Dummy for retiring

Independent Variable	Education	All Invididuals		Men		Women	
(1) 35th Highest Year > 0	All	-0.001	(0.006)	-0.006	(0.009)	0.002	(0.008)
(2) 35th Highest Year > 20% mean	All	0.013**	(0.006)	0.011	(0.007)	0.015	(0.011)
(3) 35th Highest Year > 50% mean	All	0.031***	(0.007)	0.026***	(0.007)	0.046***	(0.017)
(4) 35th Highest Year / Last 5 Years' Avg. Earnings	All	-0.014	(0.011)	-0.012	(0.014)	0.000	(0.022)
(5) 35th Highest Year > 0	Low	-0.002	(0.008)	-0.008	(0.012)	-0.002	(0.010)
(6) 35th Highest Year > 0	High	0.003	(0.009)	0.003	(0.014)	0.005	(0.012)
(7) 35th Highest Year > 20% mean	Low	0.013	(0.008)	0.015	(0.011)	0.007	(0.014)
(8) 35th Highest Year > 20% mean	High	0.015	(0.009)	0.011	(0.007)	0.022	(0.011)
(9) 35th Highest Year > 50% mean	Low	0.034***	(0.009)	0.023**	(0.010)	0.078***	(0.025)
(10) 35th Highest Year > 50% mean	High	0.027***	(0.007)	0.031***	(0.007)	0.011	(0.017)
(11) 35th Highest Year / Last 5 Years' Avg. Earnings	Low	-0.013	(0.015)	-0.010	(0.018)	0.006	(0.029)
(12) 35th Highest Year / Last 5 Years' Avg. Earnings	High	-0.026	(0.018)	-0.028	(0.023)	-0.024	(0.033)

Notes: Standard errors are in parentheses and clustered by individual.
Low education indicates high school dropout or high school graduate only. High education indicates individual attended college.
Sample restricted to non-retired individuals aged 55+ with at least 20 years' work experience and at least $5000 earned income in past 5 years.
Sample includes 4477 individuals (26195 person-year obs.), of which 2337 are male (13784 person-year obs.) and 2140 are female (12441 person-year obs.).
Of the 2337 men, 1396 have low education (7935 person-year obs.) and 941 have high education (5849 person-year obs.).
Of the 2140 women, 1315 have low education (7498 person-year obs.) and 825 have high education (4913 person-year obs.).
* indicates p-value < .10
** indicates p-value < .05
*** indicates p-value < .01

Table IV--Cox Hazard Model of Retirement

Dependent variable: Retirement

Independent Variable	Education	All Invididuals		Men		Women	
(1) 35th Highest Year > 0	All	-0.001	(0.058)	-0.065	(0.098)	0.042	(0.072)
(2) 35th Highest Year > 20% mean	All	0.086	(0.058)	0.112	(0.083)	0.097	(0.084)
(3) 35th Highest Year > 50% mean	All	0.155***	(0.060)	0.148**	(0.074)	0.222*	(0.117)
(4) 35th Highest Year / Last 5 Years' Avg. Earnings	All	-0.129	(0.095)	-0.083	(0.122)	-0.050	(0.170)
(5) 35th Highest Year > 0	Low	-0.007	(0.073)	-0.113	(0.124)	0.022	(0.091)
(6) 35th Highest Year > 0	High	0.029	(0.094)	0.014	(0.158)	0.060	(0.120)
(7) 35th Highest Year > 20% mean	Low	0.084	(0.074)	0.083	(0.127)	0.136	(0.145)
(8) 35th Highest Year > 20% mean	High	0.090	(0.094)	0.141	(0.112)	0.041	(0.106)
(9) 35th Highest Year > 50% mean	Low	0.164**	(0.078)	0.122	(0.095)	0.370**	(0.145)
(10) 35th Highest Year > 50% mean	High	0.145	(0.097)	0.208*	(0.120)	-0.042	(0.190)
(11) 35th Highest Year / Last 5 Years' Avg. Earnings	Low	-0.118	(0.119)	-0.082	(0.152)	-0.011	(0.212)
(12) 35th Highest Year / Last 5 Years' Avg. Earnings	High	-0.238	(0.169)	-0.207	(0.223)	-0.282	(0.295)

Notes: Standard errors are in parentheses and clustered by individual.
Low education indicates high school dropout or high school graduate only. High education indicates individual attended college.
Sample restricted to non-retired individuals aged 55+ with at least 20 years' work experience and at least $5000 earned income in past 5 years.
Sample includes 4477 individuals (26195 person-year obs.), of which 2337 are male (13784 person-year obs.) and 2140 are female (12441 person-year obs.).
Of the 2337 men, 1396 have low education (7935 person-year obs.) and 941 have high education (5849 person-year obs.).
Of the 2140 women, 1315 have low education (7498 person-year obs.) and 825 have high education (4913 person-year obs.).
* indicates p-value < .10
** indicates p-value < .05
*** indicates p-value < .01