Labor Supply Responses to Marginal Social Security Benefits: Evidence from Discontinuities[*]

Jeffrey B. Liebman

Erzo F.P. Luttmer

David G. Seif

December 9, 2008

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 estimate the effects of the marginal Social Security benefits that accrue with additional earnings on three measures of labor supply: retirement, hours, and labor earnings. We develop a new approach to identifying these incentive effects by exploiting five provisions in the Social Security benefit rules that generate discontinuities in marginal benefits or non-linearities in marginal benefits that converge to discontinuities as uncertainty about the future is resolved. We find clear evidence that individuals approaching retirement (age 52 and older) respond to the Social Security tax-benefit link on the extensive margin of their labor supply decisions: we estimate that a 10 percent increase in the net-of-tax share reduces the two-year retirement hazard by a statistically significant 2.1 percentage points from a base rate of 15 percent. The evidence with regards to labor supply responses on the intensive margin is more mixed: we estimate that the elasticity of hours with respect to the net-of-tax share is 0.41 and statistically significant, but we do not find a statistically significant earnings elasticity.

1. Introduction

A common argument is that investment-based Social Security reform will improve economic efficiency by increasing the perceived link between retirement contributions and retirement benefits (Auerbach and Kotlikoff, 1987; Kotlikoff, 1996; Feldstein and Liebman, 2002). Under this argument, individuals currently perceive the Old-Age, Survivors, and Disability Insurance (OASDI) payroll tax as a pure tax, failing 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. A notional defined-contribution system could similarly produce efficiency gains by making the tax-benefit link more transparent.

Though economists have long recognized Social Security's tax-benefit link (Browning, 1975; Blinder et al., 1980; Burkhauser and Turner, 1985), there is little evidence as to whether people perceive the Social Security tax as a pure tax or whether they instead 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 effective Social Security tax rate affects labor supply as measured by hours or earnings. While there is an extensive literature analyzing 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 effects of the level of Social Security Wealth). Moreover, as we explain later, nearly all of the papers that do account for accrual confound the retirement incentives with the benefit claiming date incentives. We instead isolate the retirement labor supply incentives. We see this, together with our examination of labor supply responses on the intensive-margin (hours and earnings), as the first major contribution of this paper.

A challenge that faces all research on the incentive effects of Social Security is the concern that variation in these incentives may be correlated with unobserved determinants of labor supply. Structural models explicitly exclude such unobserved determinants from the utility function and instead focus on the question of whether the resulting preferences in combination with the Social Security rules can explain observed

1

retirement patterns (Gustman and Steinmeier, 1986, 2005a; Rust and Phelan, 1997; Laitner and Silverman, 2008). Research that exploits variation over time in the Social Security rules can deal with this concern by using sharp variation in the generosity of benefits that applies to certain cohorts, as Krueger and Pischke (1992) did when using the variation generated by the "notch generation."[1] Most research that uses cross-sectional variation in incentives attempts to address the concern by including determinants of these incentives as control variables. This approach has become feasible since the early 1980s when datasets were first matched with administrative Social Security earnings histories. Such matched data were used in papers by Fields and Mitchell (1984), Burtless and Moffitt (1984), Hausman and Wise (1985), Burtless (1986), Sueyoshi (1989), McCarty (1990), Vistnes (1994), and Blau (1997). If *all* determinants of the incentives are included as controls, as is done in Coile (2004) and Coile and Gruber (2007) but not in the earlier papers, the resulting estimates will be identified off of the non-linearities in the incentive schedule that are not absorbed by the control variables. The estimates will be unbiased if unobserved determinants of labor supply are uncorrelated with these non-linearities. This is more likely when the non-linearities are strong and vary across individuals, as is the case with Samwick's (1998) variation in specific pension plan features across individuals in different firms. As explained in more detail below, we develop a methodology in which the estimated incentive effects are identified only off of provisions in Social Security benefit rules that generate discontinuities in incentives. By relying on this variation, we substantially reduce the scope for bias in our estimates from unobserved determinants of labor supply that are correlated with general non-linearities in the Social Security benefit rules. We see this methodology as the second major contribution of this paper.

The Social Security benefit formula contains a number of provisions that can create large variations in the effective marginal tax rate for otherwise very similar individuals (Boskin et al., 1987; Feldstein and Samwick, 1992). In particular, we exploit discontinuities generated by five provisions of the Social Security benefit formula. First,

---

[1] While there has been little sharp variation over time in Social Security benefit rules in the U.S., other countries have made changes in the public pension system that creates effective variation in incentive and income effects across cohorts and years. Manoli et al. (2008) use such variation in the case of Austria to identify the incentive and income effects of the public pension system on retirement decisions.

Social Security benefits depend on only 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. Second, an individual receives total benefits that are the greater of either 100 percent of the person's own retired worker benefits or 50 percent of the benefit of the individual's spouse, thus creating a discontinuity in marginal benefits around the point where the Social Security benefit of one spouse is double that of the other spouse. Third, the provisions governing Social Security benefits for widows and widowers create discontinuities in marginal benefits. Fourth, kink points in the Social Security benefit schedule create discontinuities in marginal benefits, and, fifth, there is a discontinuity at the point where the individual reaches sufficient quarters of earnings (generally 40, but lower for earlier cohorts) to become vested.

Together, these five provisions potentially create sharp discontinuities in the effective Social Security tax rate when there is no uncertainty about the future labor supply of the individual and his or her spouse. When there is still uncertainty about future labor supply, these provisions can create non-linearities that converge to discontinuities as the uncertainty gets resolved. We use the term "discontinuities-in-the-limit" to refer both to actual discontinuities and to non-linearities that converge to discontinuities. We develop a variant of the standard regression discontinuity approach so that the effects of the Social Security benefit rules on labor supply are identified off of the variation created by these discontinuities-in-the-limit. Our regressions include linear controls for *all* variables that determine the marginal Social Security tax rate, as well as many interactions and higher-order terms of these variables. We develop a criterion that determines how flexible these controls need to be in order to preserve sufficient variation due to discontinuities-in-the-limit but absorb virtually all other variation. Since the variation from the discontinuities-in-the-limit identifies our estimates, these estimates would be biased only in the unlikely case that unobserved determinants of labor supply are discontinuous or exhibit strong non-linearities at exactly the same points as the ones created by these five provisions in the Social Security benefit rules. We therefore believe it is reasonable to consider our estimates as measuring the causal effects of marginal Social Security benefits.

We perform our estimation using observations from the original cohort of the Health and Retirement Study (HRS).[2] The HRS is a longitudinal survey of individuals, and the original cohort includes people 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 the administrative Social Security earnings records of HRS sample members. We find clear evidence that individuals respond to the Social Security tax-benefit link on the extensive margin of their labor supply decisions: we estimate that a 10 percent increase in the net-of-tax share reduces the two-year retirement hazard by a statistically significant 2.1 percentage points from a base rate of 15 percent. The evidence with regard to labor supply responses on the intensive margin is more mixed: we estimate that the elasticity of hours with respect to the net-of-tax share is 0.41 and statistically significant. We do not find a statistically significant earnings elasticity, though point estimates do suggest a positive elasticity, as well.

Qualitatively, and in terms of statistical significance, the extensive-margin labor supply responses are quite robust to changes in specification, but the magnitude of the point estimates varies somewhat across specifications. The intensive-margin labor supply responses are more sensitive to changes in specification. Though we lack statistical power to estimate results within subsamples precisely, the retirement response appears to be driven mostly by the female subsample, while the hours response appears to come from the male subsample. Overall, our results clearly allow us to reject the notion that labor supply is completely unaffected by the tax-benefit link in Social Security. Our estimates, however, are not sufficiently precise to determine the exact degree to which individuals perceive this tax-benefit link. They are consistent both with a complete perception of the tax-benefit link and with only a very small fraction of the tax-benefit link being perceived.

The rest of this paper proceeds as follows: In section 2, we explain the provisions in the Social Security benefit rules that give rise to discontinuities-in-the-limit and develop a methodology that exploits variation from these discontinuities-in-the-limit.

---

[2] The HRS is sponsored by the National Institute of Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. We use the RAND HRS Version F Data file (2008).

Section 3 explains the data and our empirical specifications. Section 4 presents the results, and section 5 concludes.

2. Methodology

2.1 Brief description of the Social Security benefit rules

Social Security retirement benefits in the U.S. are based on a worker's lifetime earnings record. Each year of earnings during a worker's career is indexed to the wage level of the year the worker turns 60 by multiplying the earnings by the ratio of average earnings in the year the worker turns 60 to the average earnings in the year in which the earnings were earned. Earnings after age 60 are not indexed. A worker's average indexed monthly earnings (AIME) are calculated by summing the indexed earnings from the worker's highest 35 years of indexed earnings (including zeros if the worker worked for fewer than 35 years) and then dividing by 420 (35 × 12). Only earnings up to the maximum taxable earnings level (currently $102,000) are included in the calculations. A progressive benefit formula is then applied to determine the worker's primary insurance amount (PIA). This benefit formula replaces 90 percent of average earnings over an initial segment, 32 percent over a second segment, and 15 percent of earnings over a final segment.

The PIA is the monthly benefit a worker receives if he or she retires at the full retirement age (FRA) and claims benefits as a retired worker. The PIA is indexed for inflation. Workers may claim benefits as early as age 62, with a permanent reduction in benefits of about 6 2/3 percent per year prior to the FRA. Workers who delay claiming beyond the FRA receive increased benefits from the delayed retirement credit for each year they postpone claiming. However, delays in claiming beyond age 70 do not result in increased benefits. In married couples, the lower-earning individual receives a benefit that is the greater of his or her own benefit or 50 percent of the benefit of the higher-earnings spouse. Widows and widowers receive benefits equal to the maximum of their own benefits and the full benefits of their deceased spouses.

2.2 Sources of discontinuities-in-the-limit in marginal Social Security benefits

We identified twelve provisions in the Social Security rules that generate discontinuities-in-the-limit. Because some of these provisions depend on variables not recorded in our data set or apply to relatively few individuals, we are left with five provisions that generate the variation we exploit in our empirical analysis.[3]

First, we exploit the fact that Social Security benefits depend on only the 35 highest years of indexed earnings (the "35-year rule"). 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. If this additional year is not among the 35 highest years, then there is no marginal increase in benefits from additional work. Moreover, if there is some chance, given uncertainty about future earnings, that the additional year will no longer be among the 35 highest years of earnings at the point the person's Social Security benefits are calculated, then the 35-year-rule reduces the marginal returns to work. If the additional year of earnings is among the 35 highest, then the *average* returns from working the additional year are greater, the lower were the earnings in the replaced year. However, the *marginal* returns to working an additional hour are not affected by the level of earnings in the replaced year because, on the margin, additional earnings do not displace prior earnings.

Second, the rules on spousal benefits create variation in the effective Social Security tax rate. This variation consists of non-linearities that converge to discontinuities as uncertainty about future own and spousal labor supply gets resolved. An individual

---

[3] Social Security discontinuities not studied in this paper include: (1) Income taxation of benefits – The 1993 Omnibus Budget Reconciliation Act increased the fraction of Social Security benefits subject to income taxation for higher-income individuals, thus increasing effective Social Security tax rates for those individuals. (2) Divorce – Eligibility for spousal benefits upon divorce is limited to individuals who were married for at least 10 years, thus creating a discontinuity in marginal Social Security benefits at 10 years of marriage for individuals who might claim spousal or widow benefits. (3) Remarriage – Individuals lose eligibility for spousal benefits based on an ex-spouse upon remarriage prior to age 60, thus creating jumps in marginal Social Security benefits upon remarriage for the subgroup of individuals who would have claimed benefits based on an ex-spouse's earnings. (4) The Windfall Elimination Provision – This provision places workers who receive a government pension from a job in a sector not covered by Social Security on a different benefit schedule. (5) Changes in state "double-dipping" laws – These laws prevent workers from receiving state pensions from SS-ineligible government work if they are claiming 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 similar lifetime earnings but with different year-by-year earnings histories. (7) Children's benefits – Minor children of retirees are eligible to receive 50 percent of the retiree's benefits, which creates variation in effective marginal Social Security benefits based on the age difference between the parent and child.

receives total benefits that are the greater of 100 percent of the person's own retired worker benefit or 50 percent of the benefit of the individual's spouse. When benefits are calculated, this creates a discontinuity at the point where the ratio of own to spousal PIA equals 0.5 because individuals will claim benefits on the spousal record when the ratio falls below 0.5. In this case, there is no link on the margin between own labor earnings and Social Security benefits. A similar discontinuity occurs when the PIA ratio reaches 2.0 because, at this point, the individual's spouse will also claim benefits on the individual's earnings record. When this occurs, it will discontinuously increase the tax-benefit linkage on the margin by about 50 percent.[4]

Third, there is variation due to rules regarding widow or widower benefits. An individual with a living spouse receives the maximum of her own Social Security benefit or 50 percent of her spouse's benefit, while someone with a deceased spouse receives the maximum of her own benefit and 100 percent of her deceased spouse's benefit.[5] Thus, individuals with a living spouse will claim their own benefits in a future year with the probability that their own benefits exceed 50 percent of their spouse's benefits and the spouse is alive in that year, plus the probability that their own benefits exceed 100 percent of their spouse's benefits and the spouse is deceased in that year. Thus, even for those with a living spouse, the marginal returns to work drop discontinuously if the ratio of own to spousal PIA falls below one because this severs the link between work and the value of benefits received if widowed. Of course, any uncertainty about future own and spousal labor supply will generate uncertainty about the value of the PIA ratio at the time of benefit claiming, turning the discontinuity into a non-linearity in the return to work around the earnings level where the PIA ratio equals one.

Fourth, the AIME-PIA conversion schedule contains three segments. In the first segment the PIA increases by $0.90 for every dollar increase in the AIME, in the second segment this figure is $0.32, and in the third segment it is $0.15. These kinks in the AIME-PIA conversion schedule create two discontinuities in the returns to work. First,

---

[4] The increase is exactly 50 percent if the individual and the spouse are the same age, have the same life-expectancy and retire at the FRA. In other cases, differences in life expectancy and early retirement adjustments or delayed retirement credits can cause this increase to be somewhat larger or smaller than 50 percent.
[5] As with spousal benefits, early retirement adjustments or delayed retirement credits may result in slightly different values.

the marginal returns to work are (90-32)/90=64 percent lower for those who end up on the second segment rather than on the first segment of this schedule. Second, the marginal returns to work are (32-15)/32=53 percent lower for those ending up on the third segment rather than on the second one. For those who still face uncertainty about which segment they will be on, the returns to work will be a weighted average of the returns to work at each of the segments, weighted by the probabilities of ending up on each. This uncertainty about future earnings turns the discontinuities into non-linearities in the returns to work around the earnings levels that lead the expected AIME to cross the kink points.

Fifth, individuals need a certain number of quarters of earnings (generally 40) to qualify for benefits. This rule reduces the returns to work for earnings generated before this vesting limit is reached by the probability that this limit will still not be reached by the time the person claims benefits.

These five sources of discontinuities interact in multiple ways. For example, the 35-year rule and the vesting rule do not generate variation in the effective marginal Social Security tax rate for someone who will claim spousal benefits. Similarly, the discontinuity due to widow benefits will create a greater jump in the effective marginal Social Security tax rate for someone who is on the 32 percent segment of the AIME-PIA schedule than for someone on the 15 percent segment of this schedule. Our methodology also exploits the variation in the effective marginal Social Security tax rates generated by interactions among the five provisions.

2.3 A methodology to exploit discontinuities-in-the-limit

If individuals had perfect foresight, we could use a standard regression discontinuity design to exploit the discontinuities generated by the five provisions in the Social Security benefit rules we discussed above (e.g., see Hahn et al. 2001 for the standard regression discontinuity design). In particular, we could calculate the present discounted value of all future Social Security benefit payments for person *i* and his or her spouse: $SSW_{it}(X_{i,t-1}, X^+_{i,t-1})$, where $(X_{i,t-1}, X^+_{i,t-1})$ is the vector of individual characteristics (including own and spousal earnings) that determine Social Security benefit payments. This vector consists of a component, $X_{i,t-1}$, that is known at time *t*-1, and a component,

$X^+_{i,t-1}$, that is not yet known at that time (except under perfect foresight). The person would face an effective Social Security tax of:

$$(1) \qquad \tau^{effective}_{it}(X_{i,t-1}, X^+_{i,t-1}) = \tau^{nominal}_t - \partial SSW_{it}(X_{i,t-1}, X^+_{i,t-1}) / \partial y_{it},$$

where the derivative of *SSW* with respect to current income, y_{it}, would be evaluated at the predicted value of current income (based on past income) to avoid a mechanical relationship between current labor supply decisions and the effective tax rate. We could then run a standard regression discontinuity specification to estimate the effects of the marginal tax rate on a measure of labor supply, h_{it}:

$$(2)\ h_{it} = \alpha\left(1 - \tau^{effective}_{it}(X_{i,t-1}, X^+_{i,t-1})\right) + f(X_{i,t-1}, X^+_{i,t-1}, \boldsymbol{\beta}) + Z_{it}\boldsymbol{\gamma} + \varepsilon_{it},$$

where Z_{it} is a vector of other explanatory variables for labor supply, while α, $\boldsymbol{\beta}$, and γ are parameters to be estimated, and ε is an error term. The functional form of the net-of-tax share, $1 - \tau^{effective}_{it}$, is determined by the Social Security benefit formula and, critically, contains discontinuities. By contrast, the function $f(.)$ is a *continuous* but flexible function of exactly the same characteristics that determine the net-of-tax share. If $f(.)$ is sufficiently flexible, then α, the labor supply response to the Social Security net-of-tax share, would be identified exclusively by the discontinuities in the net-of-tax share.

In reality, of course, some of the determinants of Social Security benefits are not yet known at the time when the labor supply decision is made. We therefore estimate the labor supply response to the expected net-of-tax share by:

$$(3)\ h_{it} = \alpha\left(1 - E[\tau^{effective}_{it} \mid X_{i,t-1}]\right) + f(X_{i,t-1}, \boldsymbol{\beta}) + Z_{it}\boldsymbol{\gamma} + \varepsilon_{it},$$

Due to the expectation operator, E[.], many discontinuities in the effective marginal tax rate turn into non-linearities. These non-linearities would be fully absorbed by $f(.)$ if we were to allow $f(.)$ to be an arbitrarily flexible function of $X_{i,t-1,}$ and, as a result, the labor supply response to the net-of-tax share would no longer be identified. This creates a

dilemma. On the one hand, we want $f(.)$ to be sufficiently flexible to capture any relation between past determinants of the expected effective Social Security tax rate ($X_{i,t-1}$) and unobserved determinants of labor supply (ε_{it}). On the other hand, we require sufficient remaining variation in the effective marginal tax rate to identify the labor supply effects. The key to our methodology is the creation of a criterion that allows us to determine whether the control function $f(.)$ is sufficiently flexible.

To determine the flexibility needed in $f(.)$, we first calculate the effective marginal Social Security tax under a hypothetical set of Social Security rules that have been stripped of the provisions that create discontinuities. We refer to the Social Security rules stripped of these provisions as the "smoothed" Social Security benefit rules. In particular, we (i) eliminate the 35-year rule by letting the smoothed AIME be equal to the sum of all indexed earnings (rather than the sum of the 35 highest years of indexed earnings) divided by 35, (ii) assume, instead of the rules on spousal and widow/widower benefits, that each individual receives a fixed percentage of the benefits based on the own record and a fixed percentage of the benefits of the spousal record, where these percentages are given by the actual percentages received on average by people in our data set that have the same sex, own work/retirement status, marital status, and spousal work/retirement status, (iii) replace the kinked AIME-PIA schedule by the best-fitting quadratic schedule, and (iv) eliminate the vesting rule. The resulting "smoothed" Social Security rules closely resemble the actual rules, except that they no longer contain discontinuities.

Next, we use these smoothed rules to calculate a smoothed expected effective Social Security tax rate ($\tau_{it}^{Smoothed}$) using exactly the same method that we used to calculate the actual expected effective Social Security tax rate from the actual Social Security benefit rules. We then run auxiliary regressions of the form:

$$(4) \quad h_{it} = \alpha\left(1 - \mathrm{E}[\tau_{it}^{Smoothed} \mid X_{i,t-1}]\right) + f(X_{i,t-1}, \boldsymbol{\beta}) + Z_{it}\boldsymbol{\gamma} + \varepsilon_{it},$$

In these regressions, the effect of the smoothed effective tax rate on labor supply is purely identified off of non-linearities in the Social Security benefit schedule such as the progressive nature of the AIME-PIA schedule (now modeled as a quadratic

relationship) or the fact that the present discounted value of benefits increases as individuals age (since older individuals are closer to receiving benefit payments than younger people are). Even though some of this variation may be valid, we are not comfortable using it because many of these non-linearities may be gradual and could plausibly be correlated with unobserved determinants of labor supply. To ensure that none of this variation drives our main estimates (from equation 3), we increase the flexibility of the functional form of the control function $f(.)$ until the estimate of α in the auxiliary regressions (equation 4) becomes completely insignificant. We then use that functional form for the control function in the main regression.

This approach ensures that the estimate of the effect of the effective marginal Social Security tax rate on labor supply (as estimated by equation 3) is driven by the variation in effective tax rates from the five provisions in the Social Security rules described in section 2.2. These provisions create discontinuities-in-the-limit that are specific in the sense that they appear at particular earnings levels (e.g. at earnings such that PIA ratios reach 0.5, 1.0 or 2.0). Since unobserved determinants of labor supply are unlikely to be discontinuous or exhibit strong non-linearities at exactly the same points as the ones created by these five provisions in the Social Security benefit rules, we think it is reasonable to treat the resulting estimates as causal.

In interpreting our estimates of the coefficient α, it is worth noting that workers may make labor supply decisions over multi-year horizons and substitute hours intertemporally. For example, the 35-year rule may cause workers to avoid working a 36th year while simultaneously increasing their earnings in each of the 35 prior years. The coefficient α is therefore a combination of static responses by individuals with short planning horizons and intertemporal shifting by those with longer horizons.

3. Data and Empirical Implementation

3.1 Data

We perform our estimation using data from the original cohort of the Health and Retirement Study (HRS), a longitudinal survey that can be linked to Social Security earnings 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 percent born between 1928 and 1947). Individuals were first interviewed in 1992 and have been re-interviewed every two years. Our data extend through the seventh wave of the HRS, which was conducted in 2004. In total, the original cohort of the HRS includes 12,582 individuals who were interviewed at least once.

Key to our analysis is the fact that we have historical Social Security earnings records for most members of the original cohort of the HRS and their spouses. These records include yearly earnings (up to the Social Security contribution ceiling) from 1951 through 1991.[6] In addition, the HRS contains self-reported earnings for odd-numbered years beginning in 1991, which allows us to extend our calculations of expected Social Security Wealth beyond 1991 to each survey date.

We use several variables from the HRS to construct a measure of retirement. The HRS measures contemporaneous self-reported retirement status at each survey date, as well as the year and month that each individual retired (if the individual reports being retired). In some cases, however, individuals report being retired but nevertheless report substantial labor earnings after their retirement date. We therefore define a worker as retired if the worker says he or she is "fully retired" and if his or her earnings are below $2500.[7] For details on the exact construction of the retirement status variable as well as precise definitions of all other variables, see Appendix 1.

The HRS survey data also contain the two other dependent variables for our regressions: earnings and hours worked per week. The first of these is self-reported, with answers corresponding to the previous year. Our hours worked variable is the sum of the usual hours per week individuals report working on their primary and secondary job measured at the time of the survey. In addition, the HRS contains necessary control variables for our analysis, including age, sex, education, race, industry and occupation of the longest job held, Census region of residence, and total household wealth. Data are collected semi-annually in even years, but financial variables other than household wealth correspond to the year prior to the survey year.

---

[6] Social Security benefits for individuals in our sample do not depend on earnings from years prior to 1951.
[7] As a robustness check, we construct two alternative retirement definitions. One is based solely on the earnings record, ignoring all self-reported retirement data, and the other is based solely on self-reported retirement data, ignoring all earnings data.

In constructing our analysis sample, we exclude individuals who could not be linked to administrative Social Security records themselves or whose spouse could not be linked (about one-third of potential observations). We also exclude individuals who were already retired before the initial wave of the HRS or who had very weak past labor force attachment (about 17 percent of potential observations). In addition, we exclude widowed, separated, and divorced individuals in cases for which we have insufficient information about their former spouses to calculate benefits (about 11 percent of potential observations). Furthermore, we exclude anyone who reports ever having been disabled in the HRS (about 6 percent of potential observations), as disability changes Social Security incentives in ways that we do not have sufficient information to model correctly. Other sample restrictions result in much smaller numbers of dropped observations, leaving us with a sample of 3,971 individuals (2,269 men and 1,702 women) out of the 12,582 individuals in the original HRS cohort. See Appendix Table 1 for a full list of sample selection criteria. For our analysis of hours and earnings, slightly fewer observations are included, as described in the appendix table.

We limit our sample to person-year observations on those individuals who had not yet retired as of the prior wave of the HRS. In addition, since the primary respondents in the original HRS cohort are all age 52 or older, we include spouse person-years in our analysis sample only if the spouse is 52 years or older in that year. Taking all of these restrictions into account, our sample consists of 13,902 person-year observations.

Table 1 shows summary statistics for the key variables in our data. In each (two-year) wave, an individual has approximately a 15.1 percent chance of retiring and this hazard rate does not vary significantly by sex. Conditional on working, the average male respondent works almost 42 hours per week, while the average female respondent works 35 hours per week. Mean Social Security Wealth discounted at a 3 percent real rate is $272,153. Nearly all sample members, male and female, have had sufficient earnings histories to be eligible for Social Security benefits as retired workers. Because, in constructing our sample, we dropped most of the individuals who were non-married at the time of the first wave of the HRS, 92 percent of the person-year observations in our analysis sample came from married individuals. The average age is 60 for men and 58 for

women. On average, men have had earnings in 37 prior years and women in 27 prior years.

Figure 1 shows annualized two-year retirement hazard rates by gender. 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 more than doubles from 6 percent to above 12 percent between ages 60 to 62 and then remains relatively constant thereafter.[8]

3.2 Calculating Expected Social Security Wealth

We define the effective Social Security tax rate as the nominal Social Security tax rate (10.6 percent)[9] minus the expected Social Security marginal benefit rate, where this benefit rate is defined as the marginal effect of current labor supply on expected Social Security Wealth. Thus, the calculation of Social Security Wealth is a key element of our analysis. In addition, we include Social Security Wealth as a control variable in our regressions.

For married sample members, we define Social Security Wealth as the combination of own and spousal Social Security Wealth. More specifically, it is the expected present discounted value of all payments from the Social Security Administration to the individual and his or her spouse. Future Social Security benefits are calculated using the current Social Security benefit rules, ignoring the possibility that legislative reforms will alter program rules. We implement the Social Security benefit rules exactly to the extent we have the required information, and in our implementation incorporate rules on the treatment of spousal benefits, widow benefits, benefit reductions for early retirement, the delayed retirement credit, and the vesting rule based on quarters of earnings.[10] We model the benefits workers can claim on their own earnings record and

---

[8] We do not show retirement hazards beyond the age of 70 because we have fewer than 100 observations in each age-gender cell for ages 71 or greater.

[9] We exclude the disability insurance component of OASDI, as DI benefits are not incorporated into our model. Cushing (2005) shows that for older workers the effective DI tax rate converges to the statutory rate. In our sample, it would therefore add little variation in labor supply incentives.

[10] We do not model the Special Minimum PIA because, by our calculation, it would apply to less that 0.1 percent of our observations. In addition, we do not incorporate the Windfall Elimination Provision or state "double dipping laws" because we do not have the necessary information to do this. We exclude individuals with more than 10 years of non-FICA-covered work, and thus these provisions would apply to very few of the observations that remain. In order to model them, we would need more detailed

14

any additional benefits they are entitled to based upon the record of their living or deceased spouse. We updated the benefit calculation (i) when the individual first claims benefits, (ii) when the individual first becomes eligible to claim benefits on the spousal record, (iii) when the spouse dies, or (iv) if claiming widow benefits, when the individual first becomes eligible to claim benefits on his or her own record.[11] Further details of the benefit calculation are spelled out in Appendix 2.

Future Social Security benefits are a non-linear function of (i) own year of birth, (ii) spousal year of birth, (iii) own earnings history, (iv) spousal earnings history, (v) future own earnings, (vi) future spousal earnings, (vii) year of own death, (viii) year of spousal death, (ix) year in which the individual starts claiming benefits, and (x) year in which the spouse starts claiming benefits. Year of birth and earnings history are known, but the remaining eight variables are generally stochastic.[12] Thus, future Social Security benefits are an expectation with respect to eight variables. We reduce the dimensionality of this expectation by specifying the year of benefit take-up as a function of age and year of retirement (so, conditional on age and year of retirement, year of benefit take-up is not stochastic and we do not need to take an expectation over it). In particular, we assume the individual starts claiming benefits in the year of retirement with two exceptions: (i) if the individual retires before the early retirement age, we assume that the individual starts claiming benefits at age 62 (even if widowed and eligible at age 60), and (ii) we assume those who are not retired at age 70 will nevertheless start claiming benefits then (there is never any benefit to delaying claiming benefits beyond age 70 because the delayed retirement credit does not increase after age 70). To reduce the computational burden, we further assume that retirement occurs no later than at age 80 and that death occurs no later than at age 100.

---

information than the HRS includes about current or former work for state or the federal government, as well as pension rules applicable to such work. We also do not include child benefits (payable if the retiree has own dependent children under the age of 18) in our calculation, as they, too, apply to very few individuals in our sample.

[11] The alternative of optimizing which benefits to take each year (rather than just at these four life events) would add a great deal more complexity to our calculations but would change Social Security Wealth only minimally for most individuals. Coile et al. 2002 report that fewer than 10 percent of men retiring by the age of 62 delay claiming by a year or more. Delays in claiming by a year or more are even less prevalent for those retiring after the age of 62.

[12] In some cases, some of these variables are no longer stochastic. For example, if the spouse is no longer alive, year of spousal death is not stochastic.

We model future earnings as follows: 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.[13] Finally, the probability distribution of year of death is taken from the gender-specific cohort life tables used by the Social Security Administration, adjusted for mortality differences by race and education using the estimates from Brown et al. (2002). We assume that, conditional on own and spousal age, the own and spousal year of death and retirement are independent.

3.3 The Expected Effective Social Security Tax Rate

The Social Security benefit schedule generally has different incentive effects on the extensive and intensive margins of labor supply. Following the convention in public economics, we measure the incentive effect by the log of the net-of-tax share, $\ln(1-\tau)$, where τ is the effective marginal Social Security tax. This specification has the advantage that, if the outcome variable is also specified in logs, the coefficient on $\ln(1-\tau)$ can be interpreted as a price elasticity.

To capture the incentives on the intensive margin, we define the expected effective Social Security Intensive-margin Net-of-Tax Share (INTS) for individual i in year t as:

(5) $INTS_{it} = \ln(1 - 0.106 / 1.053 + \partial SSW_{it} / \partial \hat{y}_{it})$,

where SSW_{it} denotes the individual's expected Social Security Wealth at time t, and $\hat{y}_{it}$ denotes the person's predicted pre-Social Security tax earnings for year t.[14] Because *INTS* is endogenous to the current year's earnings, we evaluate *INTS* at the predicted

---

[13] We take this approach because Coile and Gruber (2007) found that a simple method of growing earnings at a constant rate had the best predictive performance. An alternative but computationally even more intensive approach would be to generate a series of earnings trajectories for each individual, calculate incentives separately for each trajectory, and then average over all of the possible trajectories.

[14] The 10.6 percent OASI tax is based on the contract earnings, which exclude the employer's share of the tax. Thus the tax as a fraction of the pre-Social Security tax earnings is 10.6/1.053=10.1 percent.

level of earnings, which is formed by applying the age- and gender-specific earnings growth rates to the person's previous year's earnings.

To capture the incentives on the extensive margin, we calculate the *average* effective Social Security tax rate if the individual retires at the very end rather than at the very beginning of the current year.[15] We define the expected effective Social Security Extensive-margin Net-of-Tax Share (ENTS) for individual i in year t as:

$$(6) \qquad ENTS_{it} = \ln\left(1 - 0.106 / 1.053 + \left(SSW_{it}(\text{retire in } t+1) - SSW_{it}(\text{retire in } t)\right) / \hat{y}_{it}\right).$$

To ensure that the *ENTS* captures the effects of working for an additional year, rather than the effects of delaying claiming benefits by one year, we assume benefits are first claimed in year $t+1$ (or at age 62 if year $t+1$ occurs before age 62) when calculating both SSW_{it}(retire in $t+1$) and SSW_{it}(retire in t). This separation of the retirement incentives from the benefit claiming incentives is in contrast to most of the existing empirical literature on retirement incentives, a literature in which marginal incentives to an additional year of work are calculated under the assumption that when people continue working for one more year they also delay claiming for one more year.[16] While for many individuals the labor supply and claiming decisions do indeed coincide, the efficiency arguments for personal accounts or notional defined-contribution systems rely on a misperception of the link between the work decision (rather than the claiming decision) and the level of future benefits.

4. Results

4.1 Effective Social Security Net-of-Tax Shares

---

[15] We acknowledge, but do not model, the option value in the decision not to retire, as highlighted by Stock and Wise (1990). An interesting extension would be to take a peak-value approach as in Coile and Gruber (2007) since this would make it possible to account in part for the fact that returns to work in later years might affect the decision whether or not to retire in the current year.

[16] Rust and Phelan (1997) is a notable exception in which these two decisions are treated separately. Coile et al. (2002) provide an excellent analysis of the benefit take-up decision decoupled from the retirement decision.

Before estimating the labor supply response to incremental Social Security benefits, we first present our estimates of Social Security Wealth and the corresponding intensive-margin and extensive-margin net-of-tax shares. We do this for two reasons. First, the size and variation in the incentives implicit in the Social Security rules are of interest in and of themselves, because they inform how benefit rules could be restructured to reduce the size and variation of distortions. Indeed, this is the focus of a number of papers in the literature. See, for example, Feldstein and Samwick (1992), Butricia et al. (2006), Goda (2007), Sabelhaus (2007), and Goda, Shoven, and Slavov (2009). Second, we want to document the variation in the incentives. If the variation in the estimated incentives corresponds to what we would expect given the Social Security rules, we can be more confident that our calculated incentives are correct.[17]

Figure 2 shows the distribution of Social Security Wealth in our sample, which consists of non-retired men and women between the ages of 52 and 80 and is not adjusted for family size.[18] Future benefits are discounted to the present using a 3% real discount rate. Median Social Security Wealth is $269,000 while the Social Security Wealth of 90 percent of our sample ranges between $0 and $360,578. These values are in line with those found in the literature.[19] The second and third columns of Table 2 show the mean and standard deviation of Social Security Wealth by demographic subgroup. As expected, Social Security Wealth increases with work history, lifetime earnings, and education. In addition, it is higher for married individuals than for widowed or single individuals.

Figure 3 shows the distribution of the log of the effective Social Security intensive-margin net-of-tax share (INTS), as defined by equation (5). The INTS

---

[17] We also verified that our calculator of Social Security benefits yields the identical level of benefits as the ones provided by the on-line calculator of the Social Security Administration (www.ssa.gov/retire2/AnypiaApplet.html). We performed this comparison on approximately 35 hypothetical individuals or couples. However, the Social Security Administration's online calculator is limited to calculating the PIA (i.e., it does not predict lifetime benefits given expected lifespans). In addition, it does not allow variation in the retirement date of spouses, which is precisely what yields some of the more complex scenarios when calculating PIAs and Social Security Wealth.

[18] We have no valid observations older than age 80. All such individuals in the original cohort of the HRS were either retired in the first wave or were born prior to 1920, making them subject to different Social Security benefit rules.

[19] For example, Gustman et al. (1999) calculate median Social Security wealth in the HRS to be $145,000 in 1992 dollars – equivalent to $200,000 in the 2003 dollars used in our paper. We would expect Social Security wealth to continue to increase as the HRS cohort ages, so it is not surprising that our number is about 35 percent larger.

measures the incentive effect of the effective Social Security tax on an additional dollar of earnings. For those without any tax-benefit linkage (e.g. because they are sure to claim widow benefits), the effective Social Security tax is equal to the statutory tax rate of 10.6 percent, and the log of their net-of-tax share is ln(1-0.106/1.053)=-0.106. Because additional earnings can never reduce expected Social Security benefits, this is also equal to the minimum of the log of the net-of-tax share. The mean INTS is -0.037, which corresponds to a net marginal Social Security tax rate of 3.8%. Thus, on average, the effective Social Security tax is 6.8 percentage points lower than the nominal tax due to the tax-benefit linkage. However, the tax-benefit linkage varies tremendously and is highly right skewed. Whereas 20 percent of person-years have virtually no tax-benefit linkage (INTS < -0.10), the tax-benefit is sufficiently strong for 18 percent of our sample that they face an effective Social Security subsidy (INTS > 0). The latter group consists predominantly of married individuals whose spouses are highly likely to claim off of their record and who are relatively close to the retirement age (minimizing the extent to which future benefits are discounted).

The fourth and fifth columns of Table 2 show the mean and standard deviation of INTS by demographic subgroup. Work incentives are lower for women than for men because women are more likely to claim off of the records of their spouses. Among men, work incentives are stronger for those with shorter work histories and lower lifetime earnings. These effects are driven by the progressive nature of the Social Security benefit structure and the 35-year rule, giving those with lower earnings a stronger tax-benefit linkage. Among women, we find that work incentives are much stronger if their earnings are high relative to their spouses' earnings because this makes it much more likely that they will claim based on their own record. This also explains why, despite the progressive nature of the benefit schedule, work incentives are generally relatively weak for women with short earnings histories or low lifetime earnings—there is no tax-benefit link if they claim on their spouses' records.

Figure 4 shows the distribution of the log of the effective Social Security extensive-margin net-of-tax share (ENTS), as defined by equation (6). The ENTS is a measure of the incentive effect of the effective net Social Security tax on the additional earnings if the person decides to retire next year rather than in the current year, where the

additional earnings are predicted based on the person's earnings in the previous wave. The ENTS, therefore, measures the net incentive from the Social Security system from postponing retirement by one year while keeping the date of claiming Social Security benefits constant. Because additional earnings can never reduce expected Social Security benefits, the minimum value of ENTS is -0.106 and is reached for those whose effective Social Security tax is equal to the statutory rate. This occurs for 33 percent of the sample. These individuals may have no tax-benefit linkage for a variety of reasons. For example, their current predicted indexed earnings might not be among their 35 highest annual indexed earnings, they might be sure to claim benefits on their spouse's record, or they might not be vested.

At the other extreme, 0.2 percent of observations have tax-benefit linkages that are so strong that their ENTS exceeds 0.5 (i.e., they receive an effective subsidy of about 68%). All of these cases occur because one additional year of earnings will give the individual sufficient quarters of earnings to qualify for receiving Social Security benefits. For example, we have one observation whose predicted earnings of $3,467 gives this individual a total of 40 quarters of earnings, thereby qualifying this individual for $112,308 in Social Security Wealth. This translates into an effective subsidy of 3,239% percent on earnings, or an ENTS of 3.51. While these incentives are most likely real, they are clearly outliers.[20] This produces a risk that the regressions will be driven by the handful of observations that qualify for Social Security because they reach the required number of quarters of earnings. To avoid this, we topcode ENTS at 0.50, which is slightly above the highest value of ENTS achieved by someone with more than 40 quarters of earnings. We will show below to what extent our estimates depend on this topcoding. Even after topcoding, the distribution of ENTS remains right-skewed, and has a median of -0.060 and a mean of -0.054. Ninety percent of observations in our sample have an ENTS that is negative and thus face a work disincentive. Columns 6 and 7 of Table 2 show the mean and standard deviation of ENTS by demographic subgroup. Even

---

[20] Some individuals have very few quarters of Social Security earnings because they worked most of their years in a job not covered by Social Security (often state employees covered by state pension plans). Anti-double dipping laws force them to choose between their state pension and Social Security. Thus, even when such individuals qualify for Social Security, in many cases they will choose the state pension (which is generally higher). This means that their ENTS is, in effect, much lower when they reach the 40 quarters of Social Security earnings needed to qualify for Social Security.

though the INTS and the ENTS are sometimes quite different for a given individual in a particular year, the overall variation in ENTS is similar to the variation in INTS when broken down by broad population subgroups as in Table 2.

4.2 Illustrations of discontinuities-in-the-limit on incentives

Figures 5 through 8 illustrate how provisions of the Social Security benefit rules, as described in section 2.2, create discontinuities-in-the-limit in incentives for labor supply. In other words, these provisions create non-linearities in marginal benefit rates that degenerate into discontinuities as uncertainty about future own and spousal labor supply is resolved.

Figure 5 illustrates the discontinuity in the extensive-margin incentives (ENTS) created by the 35-year rule. The dashed line plots the ENTS for a hypothetical single male who has the average lifetime earnings profile and started working at age 25. We see a sudden and dramatic drop in his ENTS at age 60, the first year in which his earnings crowd out a year of positive earnings among his 35 highest annual earnings. The solid line plots the ENTS for a second hypothetical person who is the same in all respects except that he started working at age 30. For him, we find that the drop in the ENTS occurs at age 65, the first year when previous earnings start being crowded out by current earnings in the Social Security benefits formula. These trajectories of the effective Social Security tax illustrate the type of discontinuities that help identify the labor supply responses to the effective Social Security tax.

Figure 6 shows that the incentives on the intensive margin for these same two hypothetical men exhibit no discontinuities. No discontinuities occur on the intensive margin because the predicted earnings in the current year are sufficiently high so that they are among the 35 highest annual earnings. In fact, the intensive-margin incentives are virtually identical for these two individuals. Thus, a marginal increase in earnings will have the same effect on Social Security benefits regardless of whether the current year crowds out an earlier year with positive earnings or not.

The provision that allows one to choose between claiming benefits based on 100 percent of own PIA or 50 percent of spousal PIA creates discontinuities-in-the-limit in both the intensive-margin incentives (INTS) and the extensive-margin incentives

(ENTS). Consider a hypothetical 63-year-old woman who started working at age 30 and whose earnings are equal to the mean earnings profile for women. Her husband is 63 years old, retired at age 62, started working at age 25, and in each year earned x percent of the mean earnings of males his age in that year. The solid line in Figure 7 plots the INTS of this woman as a function of x, the fraction of age-specific male mean earnings that her husband earned. We see three drops in the INTS, two of which are quite sudden. The first drop occurs around x equal to 0.2 and is associated with her husband's PIA rising above 50 percent of her expected PIA, so that he chooses to claim benefits on his rather than her record, while she continues to claim benefits on her own record. The second drop occurs more gradually between x equal to 0.7 and 1.7, because over this range it becomes increasingly likely that it will be beneficial for her to claim benefits on his record if she outlives him, and at the same time, that he will choose to claim benefits on his own record should he outlive her. In other words, uncertainty about her future labor supply has turned the discontinuity associated with claiming widow/widower benefits into a non-linearity. Finally, the third drop occurs at x equal to 2.2 percent, at which point she will most likely claim benefits on his record even when both of them are alive. As a result, her tax-benefit linkage approaches zero and her effective Social Security becomes equal to the statutory rate.[21] The short dashes plot the INTS for a woman who is the same in all respects, except that her husband died at age 63 after retiring at age 62. Her INTS now only contains one region of rapid decline, starting around x equal to 0.7, where she switches from claiming on her own record to claiming on her spousal record. The decline is only a discontinuity-in-the-limit because there is still uncertainty about her ultimate PIA due to uncertainty about her retirement date. Finally, the long dashes show the INTS for a hypothetical woman who is the same, except that she has always been single. Figure 8 plots the extensive-margin incentives (ENTS) for the same three hypothetical women, and we observe a very similar pattern of drops in incentives, except that the non-linearities are true discontinuities. We observe true discontinuities in extensive-margin incentives because these incentives measure the

---

[21] Because there is some possibility that this worker will take her own benefit before retiring (if she works long enough), her linkage rate never quite reaches zero, meaning her effective tax rate never quite reaches 10.6 percent. However, if she were a somewhat lower earner (e.g. if she had earned each year only 75 percent of the mean earnings profile for women), she would never claim on her own record for sufficiently high values of x, and her tax-benefit linkage would reach zero.

benefit of retirement in the current year relative to retirement in the next year, and therefore the women's PIAs at the retirement dates considered are known. Moreover, in this example, the husband was already retired or deceased, so his PIA was also non-stochastic.

4.3 Determination of the appropriate amount of flexibility in the control function

To ensure that the identification of the effect of labor supply incentives from the Social Security rules is driven by variation due to the discontinuities-in-the-limit from the five provisions in the Social Security benefit rules, we need to select a control function $f(X_{i,t-1}, \gamma)$ that is sufficiently flexible to absorb the remaining variation in the ENTS and INTS.

Table 3a provides a first indication of the amount of variation in the incentives that is not related to discontinuities-in-the-limit. In this table, we report the R^2 and the root mean squared error (RMSE) of regressions of the *smoothed* ENTS and INTS on increasingly flexible sets of control variables. Recall that the smoothed incentives are based on the Social Security benefit rules stripped of those provisions that create discontinuities-in-the-limit. We therefore require a control function that is able to explain the vast majority of the variation in these smoothed incentives.

Our baseline control function $f(X_{i,t-1}, \beta)$ consists of a linear combination (with weights β) of 52 lags in the log of annual earnings and the same 52 earnings history variables for the individual's spouse.[22] In addition, the linear combination contains a full set of dummies for own gender $\times$ age, for spousal gender $\times$ age, for the individual's marital status (married, widowed, never married), for own and spousal education (high school dropout, high school graduate, some college, bachelor's degree or more), for own and spousal race-ethnicity (non-Hispanic black, non-Hispanic white, Hispanic, other), for the retirement status of the individual's spouse, and for the calendar year in which the observation takes place. Further, the linear combination includes a cubic polynomial in log Social Security Wealth, a cubic polynomial in the log of the present discounted value

---

[22] All earnings are topcoded at the Social Security maximum. Unless otherwise noted, we dummy out all the logs of dollar amounts for if the dollar amount is less than $1000 in 2003 dollars. We do so for all variables in this paper that are logs of dollar amounts.

of lifetime Social Security earnings, the log of the present discounted value of spousal lifetime Social Security earnings, the number of years the spouse has been retired (if retired), the age difference between the individual and the spouse, and a cubic polynomial in the number of years in which own Social Security earnings exceeded $1000, and the same cubic polynomial for the spouse. In total, $f(X_{i,t-1}, \beta)$ is a linear combination of 433 non-collinear terms.

The baseline vector of additional controls, Z_{it} consists of the log of household assets, a cubic polynomial for job tenure at current job, for veteran status, for being born in the U.S., for longest industry of employment (13 dummies), for longest occupation (17 dummies), and for the 10 Census regions. In total, Z_{it} consists of 48 terms. We do not include the marginal labor supply incentives from private pensions as a control variable because this information is missing for a large fraction of respondents. While this would potentially be a useful control variable, its effects are unlikely to be correlated with the discontinuities-in-the-limit that we use to identify our estimates. Moreover, as Coile and Gruber (2007) show, the estimated incentive effects from Social Security are very insensitive to the inclusion of pension incentive variables.

The first row of Table 3a shows the raw variation in the smoothed ENTS and INTS in the samples that will be used to run the labor supply regressions. In the sample for the retirement regressions, the standard deviation of the smoothed ENTS is 4.55 percentage points, while the standard deviation for the smoothed INTS is 3.75 percentage points in the sample for the hours regressions and 3.86 percentage points in the sample for the earnings regressions. The fifth row of Table 3a shows that the baseline controls explain about 95 percent of the variation in the smoothed incentives and reduce the RMSE of the incentives to about 0.9 to 1.1 percentage points. Rows 2 to 4 show which components of the baseline control variables contribute the most to explaining the smoothed incentives. (Appendix Table 2 contains exact definitions of the control variables in each row.) Basic demographics only explain about 25-35%, and unless either income history or higher-order terms are included, the R^2 does not rise above about 80%. Row 6 adds an additional 163 interaction and higher-order terms to the baseline regression, which increases the explanatory power to about 97 percent. Row 7 shows that

adding a further 320 interaction and higher-order terms to the specification from row 6 only produces a very slight increase in explanatory power.

Table 3b runs the same set of regressions on the true incentives, which include variation due to the various discontinuities-in-the-limit. As expected, the RMSE is higher for the true incentives than the smoothed ones because of this additional variation. Conversely, the R^2 is lower in all specifications. The baseline set of controls is only able to explain about 58 percent of the variation in the extensive-margin incentives and about 68 percent of the variation in the intensive-margin incentives. Adding further interaction and higher-order terms can increase the explanatory power somewhat, but even in row 7, about 36 percent of the variation in the ENTS and about 26 percent of the variation in the INTS can not be explained. We conclude that this unexplained variation is primarily due to the discontinuities-in-the limit resulting from the five provisions in the Social Security benefit rules we described earlier. Overall, Tables 3a and 3b indicate that the baseline set of controls eliminates almost all variation in incentives that cannot be due to these five provisions but still leaves sufficient variation in the true incentives. The specifications of rows 6 and 7 are more conservative because they increase the fraction of unwanted variation absorbed from about 93% to almost 99%, but they do this at the cost of possibly also absorbing some of the valid variation, namely non-linearities caused by the five provisions that have not yet converged to discontinuities.

Table 4 tests whether the control function passes our criterion of absorbing sufficient variation to render all estimates on the smoothed incentive variables insignificant. In this table, we show regressions of a measure of labor supply (retirement, hours, or earnings) on the smoothed incentive measure and on various specifications of the control function. We find that the baseline control function (in row 5), as well as the more extensive ones (rows 6 and 7), pass the criterion – none of the estimated incentive effects are significant. Less flexible control functions (rows 2-4) do not pass the criterion, since they do not absorb enough variation to make the estimated incentive effects insignificant.

4.4 Labor Supply Estimates based on the Discontinuity-in-the-Limit Approach

Table 5 presents our baseline estimates of the incentive effects of the Social Security benefit rules on labor supply. The incentives are measured by the extensive-margin or intensive-margin net-of-tax shares (ENTS or INTS), and we have established that, after inclusion of our baseline controls, the primary source of identifying variation in these incentives are discontinuities-in-the-limit that arise from the 5 provisions in the Social Security benefit rules described in section 2.2. Since these discontinuities-in-the-limit are driven by specific quirks in the benefit rules, we consider it highly unlikely that there could be omitted variables that are correlated with both these discontinuities-in-the-limit and with labor supply. We therefore regard our estimates as plausibly causal.

The first row of Table 5 regresses a retirement dummy on the log of the expected effective Social Security extensive-margin net-of-tax share (ENTS) as well as the 433 baseline control variables. The first column shows the estimates for the full sample, while the next two columns separate the estimates by gender. For the full sample, we find that a higher net-of-tax share has a statistically significant negative effect on the retirement probability. In particular, an increase in the net-of-tax share of 0.10 (about two standard deviations) reduces the two-year retirement hazard by about 2.1 percentage points on a base of 15.1, or by about 14 percent. Thus, the effect is not only statistically significant but also economically meaningful. The next two columns show that this estimate appears to be driven by the subsample of women. For men, we do not find a significant effect of ENTS, but the confidence interval is sufficiently wide that we cannot rule out that the retirement response for men is as large as the point estimate for the whole sample. The hypothesis that men and women react differently to the extensive-margin incentives can be rejected only at a marginal level of significance (p-value: 0.08).

The second and third rows of the table show the responses on the intensive margin. In the second row, we use the log of usual weekly hours as the dependent variable, while the dependent variable in the third row is the log of annual earnings. Since the expected effective Social Security intensive-margin net-of-tax share (INTS) is also measured in logs, the estimates can all be interpreted as labor supply elasticities. As we are interested in measuring labor supply responses along the intensive margin, we require that the respondents be in the labor market to be included in the regression. We therefore impose a requirement of at least 15 hours of labor supply per week in the hours

regression or annual earnings of at least $2500 in the earnings regression. We estimate a statistically significant labor supply elasticity of 0.41 in whole sample when labor supply is measured by hours. As columns 2 and 3 show, this estimate is driven by the subsample of men, though the confidence interval on the estimate for women is sufficiently wide that our estimate can only rule out female labor supply elasticities larger than 0.27. We reject the hypothesis that the hours elasticity is the same for men and women. When we measure labor supply by annual earnings, we do not find statistically significant elasticities. The point estimate for the whole sample, 0.10, is positive and the corresponding 95%-confidence interval, -0.34 to +0.53, is consistent with our estimate for the hours elasticity.

Overall, we believe Table 5 provides reasonably compelling evidence that there is at least *some* effect of incentives from the Social Security benefit rules on labor supply. All the estimates in Table 5 are consistent with at least small incentive effects in the expected direction, while in four of the nine specifications we can reject the hypothesis of no incentive effects at the five percent level or better. It is important to keep in mind that our estimates are driven by responses to five discontinuities in the Social Security rules. If these discontinuities are particularly salient, then the behavioral response that we find could be greater than the behavioral response to benefit rules more generally. Moreover, our estimates are local treatment effects since they represent the responses of those individuals who are close to one or more of the discontinuities. It is possible that these individuals respond differently to labor supply incentives than the general population. One therefore needs to be careful before applying our estimates of the behavioral response to labor supply incentives to other settings.

While we can reject the hypothesis that the link between earnings and future Social Security benefits is completely ignored in people's labor supply decisions, our results provide very limited guidance as to how much of this link is perceived on average. To quantify the extent to which this link is perceived, we need to compare our estimated elasticities to labor elasticities from the literature that apply to the same population and that are identified off of incentive changes that are fully perceived. While we know of no estimates that satisfy these criteria exactly, labor supply elasticities are often estimated to be around 0.1 to 0.4 (Blundell and MaCurdy, 1999; Gruber and Saez, 2002; Kopczuk,

2005). Given this range and the confidence intervals on our estimates, our estimates are consistent both with a full perception of the tax-benefit link and with a very limited perception of this link.

4.5 Robustness

Table 6 examines the sensitivity of our baseline estimates to the specification of the control variables. Row 5 reproduces the baseline estimates. Rows 1 to 4 show regressions with less extensive specifications of the control function, and as a result these estimates may be driven by variation in the incentives not due to the five provisions of the Social Security rules that generate discontinuities-in-the-limit. Indeed, we find some instances where we estimate statistically significant incentive effects in the wrong direction (such as the earnings regression in row 2), which indicates that some variation unrelated to the discontinuities-in-the-limit is indeed correlated with unobserved determinants of labor supply.

Rows 6 and 7 show that the baseline estimates are reasonably robust to making the control function substantially more flexible (the number of control variables almost doubles from the baseline specification to the specification in row 7). The estimates for the retirement regressions are extremely robust to making the control function more flexible. This is not surprising since estimates based on true discontinuities are not sensitive to the specification of the control function (as long as it is continuous) and the sharpest discontinuities occur in the extensive-margin incentives because for these incentives there is no uncertainty about future own labor supply. The hours elasticity for the whole sample remains statistically significant in row 6 but becomes only marginally statistically significant 7. The hours elasticity for men, however, retains it statistical significance in both rows. The finding that the hours elasticity is somewhat more sensitive to the control function may be due to the fact that a very flexible control function will also absorb some of the non-linearities in incentives that have not fully converged to discontinuities.

Table 7 presents a further set of robustness tests. Panels A, B, and C show the robustness checks to the retirement regression, the hours regression, and the earnings regression, respectively. The first row of each panel reproduces the baseline regression.

The second row shows that results are very robust to including the smoothed incentives as a control variable, which provides further confirmation that the estimates are driven by the discontinuities-in-the-limit. In rows 3 and 4, we use alternative retirement definitions, based exclusively on earnings (row 3) or exclusively on self-reports (row 4) rather than our baseline definition, which combines information from both sources. We find that the results are fairly robust to the choice of the retirement definition. In the Retirement regressions, we can reject the null hypothesis of zero extensive-margin reaction at least at the 10 percent level in the case of all 3 retirement definitions. With the hours regressions, we can reject the null hypothesis at the 5 percent level in the case of two of the three measures of retirement. In the case of the earnings regressions, none of the retirement definitions produce statistically significant results. The remaining rows of panel A show that the results for the retirement regressions remain similar when we use a Probit rather than a linear probability model and are robust to how we treat the right tail of the log of the extensive-margin net-of-tax share (no topcoding at all in row 6, or in row 7, topcoding them at 0.10 rather than 0.50). Rows 5 and 6 of panel B show that the hours regressions are reasonably robust to the specification of the hours cut-off for inclusion in the sample. Finally, row 7 of panel B shows that the 1.2% of observations in which the log of the intensive-margin net-of-tax share exceeds 0.10 (i.e., where the Social Security rules provide an implicit subsidy of 10% or more), contribute importantly to the magnitude of the baseline estimate of the hours regression for the whole sample. Topcoding the INTS for these observations at 0.10 decreases the point estimate by almost half and, as a result, this estimate is no longer statistically significant. The estimate in the subsample of men, however, remains significant at the 1 percent level.

5. Conclusion

Estimating how the effective marginal Social Security tax affects labor supply is challenging, as the effective tax rate is a complicated function of own and spousal characteristics (including earnings histories), which may be correlated with unobserved determinants of labor supply. In this paper, we overcome this challenge by exploiting five provisions in the Social Security benefit rules that create discontinuities or non-linearities that converge to discontinuities as uncertainty about the future is resolved. We

develop a methodology to ensure that only this credible source of variation is being used to identify the labor supply response to the effective Social Security tax rate.

Our estimates conclusively reject the notion that labor supply is completely unresponsive to the incentives generated by the Social Security benefit rules. We find reasonably robust and statistically significant evidence that individuals are more likely to retire when the effective marginal Social Security tax is high. We also find some evidence that incentives from the Social Security rules affect labor supply along the intensive margin, but this evidence is less robust. A prominent argument has been that workers fail to perceive the link between incremental Social Security taxes paid and incremental benefits received. Our estimates contradict that argument and suggest that the potential efficiency gains from increasing the transparency of the link between Social Security benefits and taxes may be smaller than is generally assumed.

Our finding that individuals react at least to some extent to the variation in Social Security incentives raises interesting questions about the mechanism. Given the complexity of the Social Security benefit rules, how is it that workers are able to perceive and respond to the Social Security incentives? If, as Gustman and Steinmeier (2005b) show, individuals are poorly informed about the level of their Social Security benefits, is it plausible they understand the incentive effects? Are the discontinuities we study salient enough that a large fraction of the population actually understands them? Is there a way for individuals to learn optimal behavior from other, more informed, individuals without understanding the rules themselves? Chan and Stevens (2008) have answered some of these questions in the context of private pension incentives. They find that only a minority of older workers is well-informed about their retirement incentives and that this minority drives the estimated effects of pension incentives on retirement. To begin answering some of these questions in the context of Social Security, two of us have started a separate project in which we are surveying older workers and recent retirees about their understanding of Social Security rules and incentives.

References

Auerbach, Alan, and Lawrence J. Kotlikoff. 1987. *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press.

Blau, David M. 1997. "Social Security and the Labor Supply of Older Married Couples," *Labour Economics*, 4(4), pp. 373-418.

Blinder, Alan S., Roger H. Gordon, and Donald E. Wise. 1980. "Reconsidering the Work Disincentive Effects of Social Security," *National Tax Journal*, 33(4), pp. 431-42.

Blundell, Richard, and Thomas MaCurdy. 1999. "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics, vol. 3*, Orley Ashenfelter and David Card, eds. Amsterdam: Elsevier, 1999.

Boskin, Michael J., Laurence J. Kotlikoff, Douglas J. Puffert, and John B. Shoven. 1987. "Social Security: A Financial Appraisal across and within Generations," *National Tax Journal*, 40(1), pp. 19-34.

Brown, Jeffrey R., Jeffrey B. Liebman, and Joshua Pollet. 2002. "Estimating Life Tables that Reflect Socioeconomic Differences in Mortality," in *The Distributional Aspects of Social Security and Social Security Reform*, Martin Feldstein and Jeffrey B. Liebman, eds. Chicago: University of Chicago Press.

Browning, Edgar K. 1975. "Labor Supply Distortions of Social Security," *Southern Economic Journal*, 42(2), pp. 243-52.

Burkhauser, Richard V., and John A. Turner. 1985. "Is the Social Security Payroll Tax a Tax?" *Public Finance Quarterly*, 13(3), pp. 253-67.

Burtless, Gary. 1986. "Social Security, Unanticipated Benefit Increases, and the Timing of Retirement," *Review of Economic Studies*, 53(5), pp. 781-805.

Burtless, Gary and Robert Moffitt. 1984. "The Effect of Social Security Benefits on the Labor Supply of the Aged," in *Retirement and Economic Behavior*, Henry Aaron and Gary Burtless, eds. Washington, D.C.: Brookings Institution.

Butrica, Barbara A., Richard W. Johnson, Karen E. Smith, and C. Eugene Steuerle. 2006. "The Implicit Tax on Work at Older Ages," *National Tax Journal*, 59(2), pp. 211-34.

Chan, Sewin, and Ann Huff Stevens. 2008. "What You Don't Know Can't Help You: Pension Knowledge and Retirement Decision-Making," *Review of Economics and Statistics*, 90(2), pp. 253-66.

Coile, Courtney. 2004. "Retirement Incentives and Couples' Retirement Decisions," *Topics in Economic Analysis and Policy*, 4, article 17.

Coile, Courtney, Peter Diamond, Jonathan Gruber, and Alain Jousten. 2002. "Delays in Claiming Social Security Benefits," *Journal of Public Economics*, 84(3), pp. 357-85.

Coile, Courtney and Jonathan Gruber. 2007. "Future Social Security Entitlements and the Retirement Decision," *Review of Economics and Statistics*, 89(2), pp. 234-46.

Cushing, Matthew J. 2005. "Net Marginal Social Security Tax Rates over the Life Cycle," *National Tax Journal*, 58(2), pp. 227-45.

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

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

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

Fields, Gary S., and Olivia S. Mitchell. 1984. "Economic Determinants of the Optimal Retirement Age: An Empirical Investigation," *Journal of Human Resources*, 19(2), pp. 245-62.

Goda, Gopi Shah. 2007. "Could Social Security Financing Be More Efficient? Evidence from Implicit Social Security Tax Rates," Unpublished Manuscript, Harvard University.

Goda, Gopi Shah, John B. Shoven, and Sita Nataraj Slavov. 2009. "Removing the Disincentives in Social Security for Long Careers," in *Social Security Policy in a Changing Environment*, David A. Wise, ed., Chicago: University of Chicago Press.

Gruber, Jonathan, and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 84(1), pp. 1-32.

Gustman, Alan L., Olivia S. Mitchell, Andrew A. Samwick, and Thomas L. Steinmeier. 1999. "Pension and Social Security Wealth in the Health and Retirement Study," in *Wealth, Work and Health, Innovations in Measurement in the Social Sciences*, James Smith and Robert Willis, eds., Ann Arbor: University of Michigan Press, 1999.

Gustman, Alan L., and Thomas L. Steinmeier. 1986. "A Structural Retirement Model," *Econometrica*, 54(3), pp. 555-84.

Gustman, Alan L., and Thomas L. Steinmeier. 2005a. "The Social Security Early Retirement Age in a Structural Model of Retirement and Wealth," *Journal of Public Economics*, 89(2-3), pp. 441-63.

Gustman, Alan L., and Thomas L. Steinmeier. 2005b. "Imperfect Knowledge of Social Security and Pensions," *Industrial Relations*, 44(2), pp. 373-97.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), pp. 201-9.

Hausman, Jerry A., and David A. Wise. 1985. "Social Security, Health Status, and Retirement," in *Pensions, Labor, and Individual Choice,* David A. Wise, ed., Chicago: University of Chicago Press.

Kopczuk, Wojciech. 2005. "Tax Bases, Tax Rates and the Elasticity of Reported Income," *Journal of Public Economics*, 89(11-12), pp. 2093-119.

Kotlikoff, Laurance J. 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. 412-37.

Laitner, John, and Dan Silverman. 2008. "Consumption, Retirement and Social Security: Evaluating the Efficiency of Reform that Encourages Longer Careers," unpublished manuscript, University of Michigan.

Manoli, Dayanand, Kathleen Mullen, and Mathis Wagner. 2008. "Pension Benefits and Retirement Decisions: Using Income and Price Elasticities to Identify and Estimate a Structural Model of Retirement," Unpublished Manuscript, University of California – Los Angeles.

McCarty, Therese A. 1990. "The Effect of Social Security on Married Women's Labor Force Participation," *National Tax Journal*, 43(1), pp. 95-110.

RAND HRS Data, Version H. 2006. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA.

Rust, John, and Christopher Phelan. 1997. "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica* 65(4), pp. 781-831.

Sabelhaus, John. 2007. "What Is the Effective Social Security Tax on Additional Years of Work?" *National Tax Journal*, 60(3), pp. 491-506.

Samwick, Andrew A. 1998. "New Evidence on Pensions, Social Security, and the Timing of Retirement," *Journal of Public Economics*, 70(2), pp. 207-36.

Stock, James H., and David A. Wise. 1990. "Pensions, the Option Value of Work, and Retirement," *Econometrica*, 58(5), pp. 1151-80.

Sueyoshi, Glenn T. 1989. "Social Security and the Determinants of Full and Partial Retirement: A Competing Risks Analysis," NBER working paper no. 3113.

Vistnes, Jessica P. 1994. "An Empirical Analysis of Married Women's Retirement Decisions," *National Tax Journal*, 48(1), pp. 135-55.

Appendix 1: Variable Definitions

This appendix contains definitions for the variables used in our regressions, except for those variables of which the definition is obvious (such as age or sex). Unless otherwise specified, all dollar amounts are in 2003 dollars.

Dependent Variables

retired — Dummy for whether respondent retired within the previous year. Our definition of retirement is explained in more detail at the end of this appendix.

hours — Self-reported usual hours worked per week on primary and second job.

earnings — Annual earnings in the year of observation.

Independent Variables

$earnings_h$ — Log of indexed Social Security-eligible earnings h years prior to the year of observation, where h takes values 1 through 52. The final year of observations occurs in 2003, and the first year for which SS income is available is 1951. For observations prior to 2003, all values corresponding to years prior to 1951 are set to zero. The underlying earnings have a range of zero to the indexed Social Security maximum in the applicable year.

spousal $earnings_h$ — Log of spousal indexed Social Security-eligible earnings h years prior to the year of observation. In the case where the respondent is not married spousal $earnings_h$ is equal to zero.

assets — Log of total family assets, where total family assets is a value taken directly from the RAND version of the HRS.

lifetime earnings — Log of the PDV of total SS-eligible earnings since 1951 (excluding the year of observation). This is calculated by taking the log of the sum of the PDVs of each prior year's SS-eligible earnings. The PDV of a year's SS-eligible earnings calculated using 3% real interest rate.

marital status dummies — These consist of three dummies for the marital status of the respondent: married, widowed, and never married (divorced individuals are excluded from our sample).

wave dummies — Set of 7 dummy variables indicating the wave of the HRS from which the observation was taken.

education dummies — Set of 4 dummy variables for the educational attainment of the respondent: high school dropout, high school graduate, some college, and a college degree.

region dummies — Set of 10 dummies for the US Census regions of residence

race dummies — Set of 4 dummies for race/ethnicity of respondent: non-Hispanic white, non-Hispanic black, Hispanic, and other.

occupation dummies — Set of 17 dummies for the longest occupation of the respondent and an eighteenth dummy corresponding to missing. The 17 occupations are available in the RAND HRS Data Documentation, Version F, page 969. The dummy variables are generated from the RAND version of the HRS variables R1JLOCC, R2JLOCC, R3JLOCC, R4JLOCC, R5JLOCC, R6JLOCC, and R7JLOCC.

industry dummies — Set of 13 dummies for the industry in which the respondent has been employed the longest and a fourteenth dummy corresponding to missing. The 13 industries are available in the RAND HRS Data Documentation, Version F, page 974. The dummy variables are generated from the RAND version of the HRS variables R1JLIND, R2JLIND, R3JLIND, R4JLIND, R5JLIND, R6JLIND, and R7JLIND.

number of years worked in total — Number of years in which respondent had indexed Social Security-eligible earnings greater than $1000.

veteran status— Dummy for respondent's veteran status. Set to zero if unknown.

born in the USA — Dummy for whether respondent was born in the United States. Set to zero if unknown.

current job tenure—Number of years respondent has been at current job. Set to zero if unknown.

Defining Retirement

Baseline Definition

The baseline definition is based on both self-reported retirement status and reported earnings. In order to be considered retired, a person must report earnings below a certain threshold *and* report being retired.

First, a person's earnings must be and remain below $2,500 according to the SSA-provided earnings data through 1990 and the self-reported, bi-annual earnings data respondents provide to the HRS thereafter. Respondents are asked in each bi-annual wave (even years from 1992-2004) what their earnings were in the preceding year (corresponding to odd years from 1991-2003). The specifics of this question and the raw variables in the RAND version of the HRS, R1IEARN, R2IEARN, R3IEARN, R4IEARN, R5IEARN, R6IEARN, and R7IEARN are available in the RAND HRS Data Documentation, Version F, page 612. In addition, our analysis uses SSA administrative earnings data from 1951 through 1990.

Second, a person must report being retired using the self-reported retirement status RAND version HRS variables, R1SAYRET, R2SAYRET, R3SAYRET,

R4SAYRET, R5SAYRET, R6SAYRET, and R7SAYRET, information on which is available in the RAND HRS Data Documentation, Version F, page 847. Respondents must call themselves "fully retired" to be considered retired, not "partially retired." Note that in some cases, after a respondent calls himself fully retired, he is not asked whether he is retired in subsequent waves. In this case, the individual's "missing" response is considered to be the same as "fully retired." A respondent who reports being fully retired but in later waves changes his status to partially retired or not retired is *not* considered to have been retired.

Since the self-reported retirement status variables cover a two-year span, we use the variables R1RETYR, R2RETYR, R3RETYR, R4RETYR, R5RETYR, R6RETYR, and R7RETYR, which indicate the self-reported year of retirement, to determine the precise year that a respondent began considering himself retired. Information on these variables is available in the RAND HRS Data Documentation, Version F, page 853. In cases where these are missing, the year immediately before the survey took place is counted as the year of retirement. (This corresponds to the midpoint of the previous survey and the current survey.)

Thus, individuals who either report non-retirement or earnings of at least $2,500 in the first wave are considered retired in the year that their income falls below $2,500 *and* they report their retirement status as "fully retired" (assuming there is no later deviation from either status). Individuals who report being fully retired and with income below $2,500 in the first wave (and who never change their retirement status nor report income of over $2,500) are considered to be retired in the first year in which their income fell below $2,500 according to SSA records.

Individuals who never have income above $2,500 in either the self-reported HRS or the SSA-provided data are excluded from the analysis.

Earnings-Based Retirement Definition

The earnings-based retirement definition relies on the same earnings criterion used to determine retirement in the baseline retirement definition.

In this method, an individual is considered to be retired if his yearly earnings fall below $2500 and remain at that level in all subsequent observations. (In addition, earnings must have been at or above $2500 at some point; otherwise, the individual is considered never to have been in the labor force and thus never to have retired.)

The year of retirement is considered to be the midpoint between the last reported year of earnings greater than or equal to $2500 and the first year earnings permanently fell below $2500. If this average falls "between" two years, the later year is used.

Self-Reported Retirement Definition

The self-reported retirement methodology uses self-reports to determine retirement. It is based closely on the methodology of Coile and Gruber (2007).

There are 21 variables that are used to determine retirement—the 14 from the baseline definition describing self-reported retirement status and self-reported year of retirement, and seven more describing self-reported labor force status. In the RAND version of the HRS, these variables are R1LBRF, R2LBRF, R3LBRF, R4LBRF, R5LBRF, R6LBRF, and R7LBRF. Information on these variables is in the RAND HRS Data Documentation, Version F, page 911.

An individual is considered retired if his self-reported labor force status is either "retired" or "partially retired" *unless* his self-reported retirement status is "not retired at all."

As in the baseline case, since the self-reported retirement variables cover two-year periods, the precise year of retirement is then determined by the self-reported date of retirement. In cases where this is missing, the year before the survey took place is counted as the year of retirement. (This corresponds to the midpoint of the previous survey and the current survey.)

Consistent with the retirement definition of Coile and Gruber, once an individual is determined to be retired, no future information—including evidence of future labor force re-entry—is considered.

Appendix 2: Calculation of Expected Social Security Wealth

The following is a detailed description of the methodology we use to calculate family Social Security Wealth, which is then used to calculate the intensive-margin net-of-tax share (INTS) and the extensive-margin net-of-tax share (ENTS).

Social Security Wealth

Our calculation of Social Security Wealth proceeds in three broad sections. First, the program calculates prospective Primary Insurance Amounts (PIA) based on actual earnings histories and potential future earnings. Second, the program calculates probabilities of various future earnings scenarios. Finally, given the PIAs and probabilities previously determined, the program calculates expected Social Security benefits.

Calculating PIAs

The program first calculates, for each respondent and given his work history, a separate PIA based on each possible retirement date. Furthermore, it calculates a Death Primary Insurance Amount (DPIA)[23] for each possible retirement date and year of death. It then does the same for the spouse, if applicable. The program assumes that retirement occurs at some point between the current age and age 80,[24] and so separate PIAs based on each possible date of retirement are considered. We assume retirement is irreversible.

Let $\mathbf{y}_{i,t-1} = \{\dots, y_{i,t-2}, y_{i,t-1}\}$ denote the vector of the individual's indexed earnings up to and including year t-1. Only earnings up to the Social Security taxable maximum are included and these are indexed for economy-wide wage growth up to age 60. (Earnings after age 60 are not indexed.) Let the vector $\mathbf{y}^+(\mathbf{y}_{i,t-1}, i, t, c)$ denote predicted future indexed earnings if the respondent remains in the labor force. Specifically, $\mathbf{y}^+$ consists of c elements corresponding to indexed earnings in year t and in future years until, but not including, year c. Elements of $\mathbf{y}^+$ are calculated by increasing the individual's earnings in year t-1 by the age- and gender-specific wage growth rates. These wage growth rates are estimated from within the HRS, by age and sex, for individuals who remain within the labor force. In addition, we index these predicted earnings according to Social Security rules and assumptions about wage growth and inflation. Thus, the vector function $\mathbf{y}^+$ depends on $\mathbf{y}_{i,t-1}$ because $y_{i,t-1}$ is used as a baseline to predict future earnings. It depends on i because the amount of indexing on a given year's wage is dependent on the year of birth of the earner and because the wage growth rates are age- and gender-specific. Finally, $\mathbf{y}^+$ depends on t, the year of observation, because years prior to t have already occurred and therefore do not appear in the prediction of future earnings.

Let PIA($\mathbf{y}_{i,t-1}$, $\mathbf{y}^+(\mathbf{y}_{i,t-1}, i, t, \min(n, j))$, B($i$)) be the PIA in year n of individual i who retires in year j, where $B(i)$ denotes the birth date of individual i. The PIA is a function of the individual's earnings profile up to the earlier of either the year to which the PIA applies (year n) or the year in which the earnings ended due to retirement (year j). Specifically, the earnings profile is used to calculate the individual's AIME, or Average Indexed Monthly Earnings. For individuals born in 1928 or later, the AIME is the

---

[23] *PIA* refers to the Life Primary Insurance Amount of a worker. There is a separate Death Primary Insurance Amount (*DPIA*) that is bequeathed to one's spouse.
[24] PIAs are also calculated for cases in which an individual dies not yet having retired.

average of the 35 highest years of indexed earnings. According to the Social Security vesting rule, the AIME is set to zero if the individual has too few years of earnings, as defined by the Social Security Administration.[25] The AIME is then converted to a PIA based on a progressive schedule. This schedule does not vary over time, but there is a separate schedule for each cohort as defined by year of birth, which explains why the PIA is a function of B(i). Finally, the nominal PIA is indexed for inflation for every year past age 62.[26] For tractability, let $P_{itnj} = (\mathbf{y}_{i,t-1}, \mathbf{y}^{+}(\mathbf{y}_{i,t-1}, i, t, \min(n, j)), B(i))$

We denote the spousal analogues of these variable by a superscript S. Thus, let the spouse's PIA by denoted by $P_{itnk}^{S} = PIA(\mathbf{y}_{i,t-1}^{S}, \mathbf{y}^{+}(\mathbf{y}_{i,t-1}^{S}, i, t, \min(n, k)), B^{S}(i))$, where k denotes the year in which the spouse retired. Thus, P_{itnk}^{S} is the living spouse PIA for which an individual is eligible based on his living spouse's record. Note that this function is the same function used to calculate P_{itnj} above, but with different inputs—those of the spouse.

Finally, let $DP_{itnkm}^{S} = DPIA(\mathbf{y}_{it}^{S}, \mathbf{y}^{+}(\mathbf{y}_{it}^{S}, i, t, \min(n, k, m)), m, B^{S}(i), B(i))$. DPIA is the PIA that an individual's spouse bequeaths to the individual upon the spouse's death in year m. For a spouse who dies over the age of 62, it is the exact same as the regular PIA. However, special rules exist in the cases of spouses who die younger than 62. In that case, the death PIA is based on fewer than 35 years of earnings and is subject to a different indexing scheme. Because of these special rules, DPIA depends on the year of spousal death, m. These rules also vary slightly based on the age of the surviving spouse, and thus the birth year of the individual, B(i), is relevant as well.

Calculating Probabilities

Next, the program calculates the probability that each death and retirement combination takes place for each individual. The probabilities of retirement are the ones we observe within the HRS, by age and sex. The probabilities of death are based on cohort life tables released by the Social Security Administration, by sex and date of birth.[27] We modify these death probabilities to account for race and educational attainment using the method described in Brown et al. (2002).

Let d_{itl} be the probability the individual i dies in period l conditional on being alive in year t. The spousal analogue of this probability is d_{itm}^{S}, where m, as before, defines the year of spousal death. Let r_{itjl} denote the probability that individual i's earnings ended in year j as a result of voluntary retirement, conditional on not having yet retired at the beginning of year t and conditional on death in year l.[28] Let u_{itl} denote the

---

[25] The minimum number of significant years of earnings is generally 10, or 40 quarters. An individual can earns up to 4 quarters per year, and the threshold for earning a "quarter" is quite low; it is $1,090 in 2009. Rules regarding the effect of disability on one's PIA are ignored in our paper, as all individuals who report disability are excluded from our analysis.

[26] Note that inflation affects the real value of the PIA in two ways. First, neither the earnings input to the PIA nor the PIA itself is indexed for inflation from age 60 to age 62, and thus those two years of inflation erode the real value of the PIA. Second, unindexed (nominal) earnings after the age of 62 are input to the AIME (if they are in the top 35 years), which is then converted to a PIA, which is then, in turn, indexed for inflation from the age of 62.

[27] We assume that one dies immediately upon turning 100, although the SSA cohort life tables do not assume this until age 120.

[28] This means that the last year with positive earnings was in year t-1.

probability that individual i's earnings ended in year l as a result of death conditional on not having retired at the beginning of year t and conditional on death in year l.

For all values of $j \leq l$, we assume r_{itjl} is uncorrelated with the year of death, l. The probability from a situation whereby one both retires and dies in the same year accrues to r_{itjl}, as retirement clearly happened first, and death was thus not the direct cause of the end of the earnings. Thus, u_{itl} excludes the scenario whereby a person retires voluntarily and then dies the same year and u_{itl} only captures the probability that one's earnings end in l and that one also dies in year l, having never retired.[29] In effect, u_{itl} captures the sum of the retirement probabilities that *would have existed* into the future if death had not occurred in year l. Given these definitions, r_{itjl}, u_{itl}, d_{itl} are related as follows:

$$1 = \left(\sum_{j=t}^{l} r_{itjl} \right) + u_{itl}, \forall l \geq t \ ,$$

$$1 = \sum_{l=t}^{\infty} \left[d_{itl} \left(\left(\sum_{j=t}^{l} r_{itjl} \right) + u_{itl} \right) \right], \text{ and}$$

$$r_{i,t,j,l} = r_{i,t,j',l} \ , \forall t \leq j, j' \ .$$

The spousal analogues of the above are r_{itkm}^{S} and u_{itm}^{S}.

Calculating Benefits

Finally, the program calculates the stream of benefits for a couple given each possible state of the world (i.e., each combination of a respondent's and spouse's dates of retirement and death), calculates the PDV of each using a 3% real discount rate, and weights each stream's PDV by the appropriate probabilities described above.

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, an individual who is a widow can claim benefits based on the maximum of his or her own PIA and his or her spouse's DPIA.

An individual's Social Security benefits are adjusted based on his or her date of birth and age at which he or she started claiming benefits. In case of spousal or widow benefits, benefits also depend on the individual's spouse's date of birth, age of benefit claiming (if the spouse ever claimed before dying), and age of death. We assume that individuals start claiming benefits at the the greater of the year in which they turn 62 (the Early Retirement Age, ERA) and year j, their year of retirement, but that no individual postpones claiming benefits beyond the year in which they turn 70. Let A_{injl} be the benefit adjustment factor in year n for individual i taking benefits based on his or her own earnings, given a birth date of B(i), a retirement date of j, and a death date of l. This factor is zero in years before the early retirement age is reached and after death, meaning no benefits are received. Note that these four inputs, n, B(i), j, and l, fully define the age

---

[29] Under certain circumstances, if a person does voluntarily retire (and take up Social Security benefits) and then die in the same year, the PIA bequeathed to that person's spouse is different from if the person had simply died.

of take-up, whether take-up ever occurs, and the adjustment factor. A_{injl} is equal to zero if take-up has not yet occurred or if n is greater than or equal to l. The actual benefits received, if the person claims benefits on his or her own record, are equal to the adjustment factor times the PIA. A_{injl} takes into account the delayed retirement credit and "early retirement penalty."[30] The delayed retirement credit increases benefits by several percent for each year claiming is delayed beyond the Full Retirement Age (FRA), and the early retirement penalty decreases benefits by several percent (the exact percentage of each varies by cohort) for every year take-up occurs before the FRA.[31]

Let SA_{injlm} be the adjustment factor in year n for an individual taking spousal benefits (benefits based on a *living* spouse's earnings). This is function of everything of which A is a function, as well as of i's spouse's date of birth, $B^S(i)$, and of i's spouse's date of death, m. SA is zero if any of the following are true in n: i's spouse has not reached the ERA ($n < 62 + B^S(i)$), i has not taken up benefits ($n < \min\{j, 70 + B(i)\}$), i has died ($n \geq l$), or i's spouse has died ($n \geq m$).[32] The actual benefits received are equal to the adjustment factor times half of the spouse's PIA.

Let WA_{injklm} be the adjustment factor in year n for someone taking widow benefits (benefits based on his or her *deceased* spouse's benefits). This is a function of everything of which SA is a function, as well as of i's spouse's retirement date, k. WA is zero if any of the following are true in t: i's spouse is still alive, i has not yet taken up benefits, or i has died.

Given everything above, the expected own Social Security benefit in year n (where $n \geq t$) is found by taking the expectation of maximum benefits in that year over all possible combinations of the own and spousal labor exit date and own and spousal date of death. It is also necessary to differentiate between the two possible reasons for exiting the labor force, voluntary retirement and death. Because there are two possibilities for each person in the couple, there are four possible "retirement combinations," each of which contributes to the expected Social Security benefit in a given year. Conditional on information known in year t (including earnings histories upto year t-1), and letting j index year of own retirement, k index year of spousal retirement, l index year of own death, and m index year of spousal death, the portion of benefits in year n due to both individuals' voluntarily retiring is $E\left[SSB_{itn}^{r,sr} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right]$, which is calculated as:

$$E_t\left[SSB_{itn}^{r,sr} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] = \sum_{j=t}^{\infty}\sum_{k=t}^{\infty}\sum_{l=j}^{\infty}\sum_{m=k}^{\infty} d_{itl} d_{itm}^S r_{itjl} r_{itkm}^S \times$$

$$\left[1(m > n)\max\left\{ A_{injl}\cdot P_{itnj}, SA_{injlm}\cdot \tfrac{1}{2}P_{itnk}^S \right\} + 1(m \leq n)\max\left\{ A_{injl}\cdot P_{itnj}, WA_{injklm}\cdot DP_{itnkm}^S \right\} \right],$$

---

where 1(.) denotes the indicator function (which equals one if the expression between parentheses is true and zero otherwise).

The portion of benefits in year n due to voluntary retirement on the individual's part but labor force exit due to death for the individual's spouse is $E_t \left[SSB_{itn}^{r,su} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right]$. Continuing to allow j, l, and m to index year of own retirement, year of own death, and year of spousal death, respectively; and setting k equal to $\varnothing$, which previously indexed year of spousal voluntary retirement, this is:

$$E_t \left[SSB_{itn}^{r,su} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] = \sum_{j=t}^{\infty} \sum_{l=j}^{\infty} \sum_{m=t}^{\infty} d_{itl} d_{itm}^S r_{itjl} u_{itm}^S \times$$
$$\left[1(m > n) \max \left\{ A_{injl} \cdot P_{itnj}, SA_{injlm} \cdot \tfrac{1}{2} P_{itnk}^S \right\} + 1(m \le n) \max \left\{ A_{injl} \cdot P_{itnj}, WA_{injklm} \cdot DP_{itnkm}^S \right\} \right].$$

The portion of benefits in n due to voluntary spousal retirement but labor force exit due to death on the individual's part is $E_t \left[SSB_{itn}^{u,sr} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^S \right]$. Calculating this sets j, year of own retirement, to $\varnothing$, which yields:

$$E_t \left[SSB_{itn}^{u,sr} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] = \sum_{k=t}^{\infty} \sum_{l=t}^{\infty} \sum_{m=k}^{\infty} d_{itl} d_{itm}^S u_{itl} r_{itkm}^S \times$$
$$\left[1(m > n) \max \left\{ A_{injl} \cdot P_{itnj}, SA_{injlm} \cdot \tfrac{1}{2} P_{itnk}^S \right\} + 1(m \le n) \max \left\{ A_{injl} \cdot P_{itnj}, WA_{injklm} \cdot DP_{itnkm}^S \right\} \right].$$

Finally, there is the portion of benefits in n due to both members of a married couple leaving the labor force due to death, which is $E_t \left[SSB_{itn}^{u,su} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right]$. It is rare, but possible, that benefits are collected under this scenario. It occurs solely when a widowed individual, whose deceased spouse worked up until death, continues working past age 70 (at which point benefits are collected even though retirement has not yet occurred) eventually dies without having retired himself. Omitting both j and k summations and setting each to $\varnothing$, this is:

$$E_t \left[SSB_{itn}^{u,su} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] = \sum_{l=t}^{\infty} \sum_{m=t}^{\infty} d_{itl} d_{itm}^S u_{itl} u_{itm}^S \times$$
$$\left[1(m > n) \max \left\{ A_{injl} \cdot P_{itnj}, SA_{injlm} \cdot \tfrac{1}{2} P_{itnk}^S \right\} + 1(m \le n) \max \left\{ A_{injl} \cdot P_{itnj}, WA_{injklm} \cdot DP_{itnkm}^S \right\} \right].$$

The total expected Social Security benefit for person i in n consists of the sum of the benefits in these four cases:

$$E_t \left[SSB_{itn} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] = E_t \left[SSB_{itn}^{r,sr} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] + E_t \left[SSB_{itn}^{r,su} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] +$$
$$E_t \left[SSB_{itn}^{u,sr} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right] + E_t \left[SSB_{itn}^{u,su} \mid \mathbf{y}_{i,t-1}, \mathbf{y}_{i,t-1}^S \right]$$

The expected Social Security benefit in year n for i's spouse, $E_t \left[SSB_{itn}^S \mid \mathbf{y}_{i,t-1}^S, \mathbf{y}_{i,t-1} \right]$, is calculated exactly the same way as is i's benefit, but with i's and i's spouse's variables switched in all places. Overall, we define a family's total Social Security Wealth (*SSW*) as the present discounted value of all years of both the individual's and the individual's spouse's Social Security benefits. This is:

$$SSW_{it}(\mathbf{y}_{i,t-1},\mathbf{y}_{i,t-1}^{s}) = \sum_{n=t}^{\infty}\left(E_{t}[SSB_{itn} \mid \mathbf{y}_{i,t-1},\mathbf{y}_{i,t-1}^{s}] + E_{t}[SSB_{itn}^{S} \mid \mathbf{y}_{i,t-1}^{s},\mathbf{y}_{i,t-1}]\right)(1+\rho)^{-(n-t)},$$

where ρ denotes the real discount rate, which is set at 3%.

Smoothed Social Security Wealth

The smoothed Social Security Wealth calculator modifies four elements of the regular calculator to eliminate the five sources of discontinuities-in-the-limit: the vesting rule based on quarters of earnings, the 35-years rule, the kinks in the AIME-PIA conversion, the choice of taking benefits from one's own or one's living spouse's record, and the treatment of widow benefits.

To eliminate the discontinuity-in-the-limit caused by the minimum 40 quarters rule, we simply allow for benefits to be received no matter how few quarters a person worked.

To remove the discontinuity-in-the-limit induced by the 35-years rule, we allow for an individual's AIME to be calculated from as many years as he works, even if that number is greater than 35. The total earnings of all years are still treated as though they occurred over just 35 years.

We develop a new AIME-PIA conversion formula to eliminate the discontinuities-in-the-limit that exist when the conversion rate changes from 90% to 32% and from 32% to 15%. We fit a quadradic formula, $PIA = a \cdot AIME^{2} + b \cdot AIME + c$, such that the following three requirements hold. Taking the AIMEs from the *regular* Social Security Wealth calculator and placing them through the *smoothed* AIME-PIA conversion, we require that the average PIA is the same, the average $\partial PIA / \partial AIME$ is the same, and the maximum AIME yields the same PIA as under the regular method.

The largest change in the calculator is the elimination of the discontinuities-in-the-limit caused by the choice between taking one's own benefits or taking benefits based on one's spouse's record (either living or deceased). First, the program sorts individuals into eight groups by sex, marital status (married, widowed, single), and whether one's spouse is retired (only applicable for married individuals).[33] It then calculates, by group, the average fraction of benefits that *are* taken off of an individual's own record compared to the total that *could be* taken. Thus, if widowed women could possibly take $10,000 on average off of their own records but, in practice, only take $1,500 on average, the fraction applied to them would be 0.15.[34] It then performs the same computation for benefits off of spousal records. Note that there is no need for these two fractions to sum to 1.

The program then proceeds to calculate benefits. However, individuals always take a portion of the benefits based off of *both* their own and their spouse's records, and the portion they take in each case is equal to the fraction calculated above.

Calculating INTS

---

[33] As the program predicts further into the future, an individual's group might change should the spouse retire and/or die.

[34] Taking the full $10,000 would require taking benefits off of their own records even when benefits off of their deceased husbands' records were higher.

To calculate the intensive-margin net-of-tax share (INTS), the program calculating Social Security Wealth is run twice. The first time, it is run normally, as described above. The second time, one change is made—the current year's earnings are increased by $1000.[35] However, the rest of the program, including all probabilities remains the same.

The dollar difference between the second and the first calculations of Social Security Wealth, divided by 1000, is our estimate of the derivative of Social Security Wealth with respect to current earnings. As described in the body of the paper, INTS is then calculated based on the following formula:

$$INTS_{it} = \ln(1 - 0.106 / 1.053 + \partial SSW_{it} / \partial \hat{y}_{it})$$

Calculating ENTS

To calculate the extensive-margin net-of-tax share (ENTS), the program calculating Social Security Wealth also runs twice, but both versions are modified—actually, simplified—from the above description. In the first calculation, the dates of retirement and take-up of benefits are assumed to be exactly one year in the future[36] (but take-up still never occurs before age 62 or after age 70).

In the second iteration, a person is assumed to retire *this year* but take up benefits *next year*. (Again, take-up still occurs no earlier than 62 nor later than 70.) This is a modification of the take-up rule described above, and it ensures that the timing of retirement between the first and second iterations is exactly one year apart while take-up of benefits in both cases occurs at the same time.

The dollar difference between the first and second calculations of Social Security Wealth, divided by the predicted earnings in the additional year,[37] is the change in Social Security Wealth with respect to an additional year of work. As described in the body of the paper, ENTS is then calculated based on the following formula:

$$ENTS_{it} = \ln\left(1 - 0.106 / 1.053 + \left(SSW_{it}(\text{retire in } t+1) - SSW_{it}(\text{retire in } t)\right) / \hat{y}_{it}\right)$$

---

[35] All future years' predicted earnings are based on the original amount *without* the extra $1000.

[36] This applies only to the respondent, *not* to the spouse.

[37] In cases where the predicted income from the extra year of work was less than $2500, $2500 was used instead. $2500 is therefore the minimum by which the difference between these two values can be divided.

Figures

Figure 1
Annualized Two-Year Retirement Hazard

Figure 2
Expected Family Social Security Wealth

**Figure 3
Social Security Net-of-Tax Share, Intensive Margin**

In addition, there are 1283 observations with the minimum value, -0.106

**Figure 4
Social Security Net-of-Tax Share, Extensive Margin**

In addition, there are 3196 observations with the minimum value, -0.106

Figure 5
Social Security Net-of-Tax Share, Extensive Margin,
by Age of Labor Force Entry

Figure 6
Social Security Net-of-Tax Share, Intensive Margin,
by Age of Labor Force Entry

48

Figure 7
Social Security Net-of-Tax Share, Intensive Margin,
Spouses of Retirees and Widows

Figure 8
Social Security Net-of-Tax Share, Extensive Margin,
Spouses of Retirees and Widows

Table 1
Selected Summary Statistics

	Sample	Entire Sample		Men		Women	
		Mean	(std.dev.)	Mean	(std.dev.)	Mean	(std.dev.)
Earnings	C	31,622	(19,580)	35,076	(20,053)	27,551	(18,186)
ln(Earnings)	C	10.123	(0.759)	10.244	(0.746)	9.979	(0.750)
Two Year Retirement Hazard	A	0.151	(0.358)	0.158	(0.364)	0.143	(0.350)
Hours Worked per Week	A	38.82	(17.58)	41.68	(18.22)	34.97	(15.89)
Dummy for Weekly Hours≥15	A	0.888	(0.315)	0.893	(0.309)	0.882	(0.322)
Weekly Hours if Weekly Hours≥15	B	42.86	(12.08)	45.95	(12.14)	39.18	(10.91)
ln(Weekly Hours) if Weekly Hours≥15	B	3.718	(0.287)	3.794	(0.264)	3.629	(0.288)
Fraction Retired at Age 60	D	0.409	(0.492)	0.305	(0.461)	0.497	(0.500)
Fraction Retired at Age 65	D	0.759	(0.428)	0.686	(0.464)	0.821	(0.383)
Fraction Retired at Age 70	D	0.939	(0.239)	0.907	(0.290)	0.966	(0.181)
ln(SS Wealth)	A	12.46	(0.35)	12.48	(0.34)	12.43	(0.37)
SS Wealth	A	272,153	(80,940)	277,704	(78,901)	264,683	(83,029)
SS Wealth if Age≥62	A	300,493	(87,377)	310,742	(83,045)	279,415	(92,185)
Years of Earnings	A	33.10	(8.93)	37.28	(6.75)	27.49	(8.41)
Eligible for SS with Own Record	A	0.988	(0.110)	0.994	(0.077)	0.979	(0.142)
Age	A	59.40	(4.63)	60.09	(4.79)	58.47	(4.25)
Married	A	0.923	(0.266)	0.955	(0.207)	0.880	(0.325)
Widowed	A	0.035	(0.185)	0.015	(0.122)	0.063	(0.242)
Single	A	0.041	(0.199)	0.030	(0.169)	0.057	(0.232)

Notes: All dollars are 2003 dollars. Eligibility for Social Security based on own record occurs with 40 quarters of positive earnings for individuals born in 1928 or later, fewer for individuals born in 1920-1927. Sample A is the sample in the retirement regressions and contains 13,902 person*year observations (7975 men and 5927 women) from 3971 unique persons (2269 men and 1702 women). Sample B is the sample in the hours regressions and contains 10,840 person*year obserations (5891 men and 4949 women) from 3152 unique persons (1731 men and 1421 women). Sample C is the sample in the earnings regressions and contains 11,062 person*year observations (5984 men and 5078 women) from 3467 unique persons (1911 men and 1556 women). Sample D is the entire HRS cohort with reliable earnings and retirement data and contains 8337 unique persons (4508 men and 3829 women).

Table 2
Effective Social Security Net-of-Tax Shares

	Number of Obs.	Social Security Wealth		Log of Intensive-margin Net-of-Tax Share (INTS)		Log of Extensive-margin Net-of-Tax Share (ENTS)	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Panel A: Whole Sample							
All Observations	13902	272,153	(80940)	-0.037	(0.045)	-0.054	(0.050)
All Men	7975	277,704	(78901)	-0.029	(0.047)	-0.048	(0.051)
All Women	5927	264,683	(78901)	-0.048	(0.039)	-0.063	(0.051)
Panel B: Men Only							
By Work History:							
35 Years of Work or Fewer	2335	224,758	(74636)	-0.008	(0.047)	-0.007	(0.060)
More than 35 Years of Work	5640	299,625	(69685)	-0.037	(0.044)	-0.064	(0.034)
By Marital Status:							
Married	7619	284,469	(73481)	-0.028	(0.047)	-0.047	(0.051)
Widowed	120	157,300	(40547)	-0.058	(0.035)	-0.073	(0.031)
Single	236	120,521	(40883)	-0.045	(0.036)	-0.049	(0.046)
By Education:							
Low Education	4446	253,283	(71258)	-0.026	(0.050)	-0.052	(0.052)
High Education	3529	308,471	(77296)	-0.032	(0.043)	-0.043	(0.048)
By Lifetime Earnings:							
Lifetime Earnings < Median	3997	234,182	(70219)	-0.008	(0.050)	-0.028	(0.059)
Lifetime Earnings ≥ Median	3978	321,435	(60918)	-0.050	(0.032)	-0.067	(0.028)
By Ratio of Own to Spousal Lifetime Earnings:							
Ratio of Earnings 1st Quartile	1905	287,731	(84527)	-0.029	(0.043)	-0.049	(0.048)
Ratio of Earnings 2nd Quartile	1905	292,073	(69358)	-0.033	(0.043)	-0.051	(0.047)
Ratio of Earnings 3rd Quartile	1905	284,912	(66126)	-0.025	(0.049)	-0.045	(0.053)
Ratio of Earnings 4th Quartile	1904	273,157	(71285)	-0.024	(0.052)	-0.045	(0.055)
Panel C: Women Only							
By Work History:							
35 Years of Work or Fewer	4869	259,852	(81536)	-0.049	(0.039)	-0.062	(0.049)
More than 35 Years of Work	1058	286,917	(86193)	-0.044	(0.038)	-0.068	(0.032)
By Marital Status:							
Married	5217	280,376	(73935)	-0.048	(0.037)	-0.064	(0.044)
Widowed	371	153,622	(38933)	-0.067	(0.048)	-0.084	(0.055)
Single	339	144,721	(55446)	-0.026	(0.043)	-0.028	(0.058)
By Education:							
Low Education	3484	245,974	(76626)	-0.052	(0.038)	-0.068	(0.044)
High Education	2443	291,365	(84515)	-0.043	(0.039)	-0.056	(0.050)
By Lifetime Earnings:							
Lifetime Earnings < Median	2971	238,529	(77313)	-0.055	(0.044)	-0.072	(0.055)
Lifetime Earnings ≥ Median	2956	290,970	(80238)	-0.041	(0.031)	-0.055	(0.034)
By Ratio of Own to Spousal Lifetime Earnings:							
Ratio of Earnings 1st Quartile	1305	274,723	(60057)	-0.079	(0.027)	-0.099	(0.024)
Ratio of Earnings 2nd Quartile	1304	286,963	(68235)	-0.050	(0.027)	-0.071	(0.037)
Ratio of Earnings 3rd Quartile	1304	291,942	(75412)	-0.038	(0.029)	-0.050	(0.032)
Ratio of Earnings 4th Quartile	1304	267,880	(86973)	-0.027	(0.040)	-0.036	(0.050)

Notes: Low education consists of high school dropouts and high school graduates; high education consists of individuals with some college or a college degree. The sample for the ratio of own to spousal earnings is limited to married individuals. Median lifetime earnings and quartiles of the ratio of own to spousal lifetime earnings are gender-specific.

Table 3a
Regressions of "Smoothed" Incentives on Control Variables

Table reports: R^2 (top number) and Root MSE (bottom number)	Sample for Retirement Regressions Dep. Variable: Smoothed ENTS			Sample for Hours Regressions Dep. Variable: Smoothed INTS			Sample for Earnings Regressions Dep. Variable: Smoothed INTS		
	Entire Sample	Men	Women	Entire Sample	Men	Women	Entire Sample	Men	Women
1. No Controls	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0455	0.0414	0.0209	0.0375	0.0416	0.0210	0.0386	0.0430	0.0216
2. Only Basic Demographics	0.2384	0.3678	0.6853	0.3306	0.4519	0.5458	0.3422	0.4570	0.5609
	0.0397	0.0330	0.0117	0.0307	0.0309	0.0142	0.0313	0.0317	0.0143
3. Baseline Minus Earnings History	0.8115	0.7996	0.8219	0.7636	0.8323	0.7812	0.7695	0.8383	0.7972
	0.0200	0.0188	0.0090	0.0185	0.0174	0.0101	0.0188	0.0176	0.0099
4. Baseline Minus Higher Order Terms	0.9255	0.9080	0.8543	0.9174	0.9364	0.8369	0.9188	0.9364	0.8528
	0.0126	0.0128	0.0082	0.0110	0.0108	0.0088	0.0112	0.0111	0.0085
5. Baseline Controls	0.9461	0.9378	0.8645	0.9487	0.9727	0.8456	0.9503	0.9720	0.8620
	0.0107	0.0106	0.0079	0.0087	0.0071	0.0086	0.0088	0.0074	0.0083
6. Baseline Plus Additional Interactions	0.9742	0.9637	0.9558	0.9740	0.9847	0.9330	0.9752	0.9843	0.9400
	0.0075	0.0081	0.0046	0.0062	0.0054	0.0057	0.0062	0.0056	0.0055
7. Baseline Plus Further Additional Interactions	0.9765	0.9670	0.9587	0.9768	0.9866	0.9379	0.9777	0.9861	0.9443
	0.0072	0.0078	0.0045	0.0059	0.0051	0.0055	0.0059	0.0053	0.0054

Notes: In each cell, the top number is the R-squared and the bottom is the root mean squared error of a regression of the smoothed ENTS (log of the effective Social Security Extensive-margin Net-Of-Tax Share) or smoothed INTS (log of the effective Social Security Intensive-margin Net-Of-Tax Share) on the set of control variables indicated in the row. Appendix Table 2 contains detailed information on the exact set of control variables included in each row.

Table 3b
Regressions of True Incentives on Control Variables

Table reports: R² (top number) and Root MSE (bottom number)	Sample for Retirement Regressions Dependent Variable: ENTS			Sample for Hours Regressions Dependent Variable:INTS			Sample for Earnings Regressions Dependent Variable: INTS		
	Entire Sample	Men	Women	Entire Sample	Men	Women	Entire Sample	Men	Women
1. No Controls	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0495	0.0505	0.0467	0.0434	0.0459	0.0359	0.0450	0.0477	0.0376
2. Only Basic Demographics	0.0852	0.0721	0.1531	0.1376	0.0607	0.2704	0.1317	0.0570	0.2498
	0.0474	0.0487	0.0430	0.0403	0.0445	0.0307	0.0420	0.0463	0.0326
3. Baseline Minus Earnings History	0.4337	0.5142	0.4701	0.5716	0.5016	0.7262	0.5757	0.5187	0.7107
	0.0377	0.0358	0.0347	0.0288	0.0330	0.0192	0.0297	0.0337	0.0207
4. Baseline Minus Higher Order Terms	0.5471	0.5907	0.5423	0.6503	0.5878	0.7710	0.6481	0.5958	0.7526
	0.0338	0.0330	0.0325	0.0262	0.0303	0.0178	0.0272	0.0311	0.0193
5. Baseline Controls	0.5752	0.6167	0.5674	0.6780	0.6164	0.7814	0.6746	0.6206	0.7652
	0.0328	0.0320	0.0316	0.0251	0.0293	0.0174	0.0261	0.0302	0.0188
6. Baseline Plus Additional Interactions	0.6285	0.6722	0.6066	0.7210	0.6688	0.8134	0.7167	0.6702	0.7982
	0.0308	0.0298	0.0305	0.0236	0.0275	0.0163	0.0245	0.0284	0.0177
7. Baseline Plus Further Additional Interactions	0.6432	0.6832	0.6272	0.7407	0.6933	0.8261	0.7349	0.6928	0.8088
	0.0304	0.0295	0.0299	0.0229	0.0267	0.0159	0.0239	0.0276	0.0173

Notes: In each cell, the top number is the R-squared and the bottom is the root mean squared error of a regression of the ENTS (log of the effective Social Security Extensive-margin Net-Of-Tax Share) or INTS (log of the effective Social Security Intensive-margin Net-Of-Tax Share) on the set of control variables indicated in the row. Appendix Table 2 contains detailed information on the exact set of control variables included in each row.

Table 4
Effects with Different Sets of Control Variables, "Smoothed" Incentives

Controls	Retirement Regressions Dependent Variable: Retirement			Hours Regressions Dependent Variable: ln(Hours)			Earnings Regressions Dependent Variable: ln(Earnings)		
	Entire Sample	Men	Women	Entire Sample	Men	Women	Entire Sample	Men	Women
1. No Controls	0.658***	0.965***	1.046***	0.353**	-0.326**	-2.651***	-6.117***	-8.461***	-12.834***
	(0.075)	(0.118)	(0.277)	(0.140)	(0.165)	(0.351)	(0.266)	(0.314)	(0.880)
2. Only Basic Demographics	-0.056	-0.207	0.076	1.70***	0.737***	-1.840***	-0.476***	-1.818***	-4.559***
	(0.088)	(0.150)	(0.446)	(0.162)	(0.211)	(0.430)	(0.176)	(0.273)	(0.678)
3. Baseline Minus Earnings History	0.066	-0.080	0.189	1.076***	0.499	-0.721	-1.014***	-1.116**	-3.081***
	(0.167)	(0.233)	(0.551)	(0.235)	(0.326)	(0.511)	(0.307)	(0.446)	(0.903)
4. Baseline Minus Higher Order Terms	0.425*	0.231	0.683	0.015	-0.239	-0.223	-0.840	-0.613	-1.049
	(0.241)	(0.319)	(0.575)	(0.383)	(0.539)	(0.555)	(0.535)	(0.775)	(0.934)
5. Baseline Controls	0.308	-0.264	0.705	0.325	0.777	-0.378	-0.594	-0.071	-0.524
	(0.283)	(0.379)	(0.593)	(0.456)	(0.800)	(0.518)	(0.678)	(1.151)	(0.981)
6. Baseline Plus Additional Interactions	0.053	-0.283	0.230	-0.325	-0.284	-0.814	-0.957	0.287	-0.988
	(0.394)	(0.483)	(1.004)	(0.561)	(0.910)	(0.759)	(0.877)	(1.525)	(1.268)
7. Baseline Plus Further Additional Interactions	0.101	-0.235	0.383	0.284	0.738	-0.364	-0.680	0.495	-0.566
	(0.407)	(0.498)	(0.975)	(0.560)	(0.907)	(0.738)	(0.894)	(1.586)	(1.347)

Notes: The table reports the coefficient and standard error on the "smoothed" INTS for ln(Earnings) and ln(Hours) regressions, and on "smoothed" ENTS for Retirement regressions. ln(Hours) regressions are limited to individuals who report working at least 15 hours per week in a typical week. Earnings regressions are limited to individuals with at least $2,500 in earnings in the sample year. All dollars are 2003 dollars. Standard errors, clustered by individual, are in parentheses. * indicates p-value<.10; ** indicates p-value<.05; *** indicates p-value<.01. Appendix Table 2 contains detailed information on the exact set of control variables included in each row.

Table 5
Baseline Specification

Dependent Variable	Entire Sample	Men	Women	p-value Men=Women
1. Retirement				
Coefficient on ENTS	-0.205**	-0.046	-0.335***	0.075*
(Standard Error)	(0.082)	(0.115)	(0.121)	
R^2	0.222	0.242	0.240	
N	13,902	7,975	5,927	
2. ln(Hours)				
Coefficient on INTS	0.414***	0.694***	-0.334	0.003***
(Standard Error)	(0.157)	(0.174)	(0.310)	
R^2	0.276	0.242	0.301	
N	10,840	5,891	4,949	
3. ln(Earnings)				
Coefficient on INTS	0.097	0.006	0.442	0.245
(Standard Error)	(0.221)	(0.264)	(0.429)	
R^2	0.512	0.515	0.525	
N	11,062	5,984	5,078	

Notes: Independent variable is INTS for ln(Earnings) and ln(Hours) regressions, ENTS for Retirement regressions. ln(Hours) regressions are limited to individuals who report working at least 15 hours per week, on average. Earnings regressions are limited to individuals with at least $2,500 in earnings in the sample year. All dollars are 2003 dollars. Standard errors, clustered by individual, are in parentheses. * indicates p-value<.10;   ** indicates p-value<.05; *** indicates p-value<.01. Final column is a t-test of the equality of the coefficient on men and women. Appendix Table 2 contains detailed information on the exact variables included in the baseline set of controls.

Table 6
Effects with Different Sets of Control Variables, True Incentives

Controls	Retirement Regressions Dependent Variable: Retirement			Hours Regressions Dependent Variable: ln(Hours)			Earnings Regressions Dependent Variable: ln(Earnings)		
	Entire Sample	Men	Women	Entire Sample	Men	Women	Entire Sample	Men	Women
1. No Controls	-0.195***	-0.361***	-0.006	1.190***	0.693***	0.740***	1.238***	-0.697*	2.620***
	(0.066)	(0.093)	(0.10)	(0.125)	(0.144)	(0.256)	(0.302)	(0.367)	(0.534)
2. Only Basic Demographics	-0.127*	-0.075	-0.222**	1.031***	0.754***	0.419	-0.396***	-1.018***	-0.089
	(0.070)	(0.092)	(0.112)	(0.136)	(0.154)	(0.264)	(0.133)	(0.172)	(0.253)
3. Baseline Minus Earnings History	-0.062	-0.034	-0.208	0.619***	0.575***	0.186	0.287	-0.260	1.705***
	(0.077)	(0.105)	(0.128)	(0.141)	(0.158)	(0.315)	(0.198)	(0.238)	(0.394)
4. Baseline Minus Higher Order Terms	-0.219***	-0.063	-0.377***	0.304*	0.563***	-0.447	0.072	-0.030	0.370
	(0.079)	(0.110)	(0.122)	(0.156)	(0.173)	(0.312)	(0.212)	(0.256)	(0.415)
5. Baseline Controls	-0.205**	-0.046	-0.335***	0.414***	0.694***	-0.334	0.097	0.006	0.442
	(0.082)	(0.115)	(0.121)	(0.157)	(0.174)	(0.310)	(0.221)	(0.264)	(0.429)
6. Baseline Plus Additional Interactions	-0.223***	0.007	-0.362***	0.315**	0.608***	-0.390	0.053	0.015	0.319
	(0.086)	(0.124)	(0.129)	(0.160)	(0.176)	(0.330)	(0.241)	(0.290)	(0.461)
7. Baseline Plus Further Additional Interactions	-0.193**	0.004	-0.320**	0.270*	0.532***	-0.287	0.035	-0.011	0.320
	(0.086)	(0.124)	(0.130)	(0.162)	(0.181)	(0.341)	(0.245)	(0.295)	(0.469)

Notes: The table reports the coefficient and standard error on the INTS for ln(Earnings) regressions, the INTS for ln(Hours) regressions, and the ENTS for Retirement regressions. ln(Hours) regressions are limited to individuals who report working at least 15 hours per week, on average. ln(Earnings) regressions are limited to individuals with at least $2,500 in earnings in the sample year. All dollars are 2003 dollars. Standard errors, clustered by individual, are in parentheses. * indicates p-value<.10; ** indicates p-value<.05; *** indicates p-value<.01. Appendix Table 2 contains detailed information on the exact set of control variables included in each row.

Table 7
Robustness Checks

	Entire Sample		Men		Women	
Panel A: Retirement Regressions (ENTS)						
1. Baseline	-0.205**	(0.082)	-0.046	(0.115)	-0.335***	(0.121)
2. Smoothed Retirement Incentive Controls	-0.243***	(0.084)	-0.024	(0.120)	-0.377***	(0.122)
3. Earnings-Based Retirement Definition	-0.135*	(0.081)	-0.031	(0.113)	-0.183	(0.122)
4. Retirement Definition Based on Self-reports	-0.209**	(0.098)	-0.303**	(0.141)	-0.180	(0.149)
5. Probit of Baseline Regression	-0.185**	(0.076)	-0.092	(0.106)	-0.282	(0.444)
6. No Windsorization of ENTS	-0.135***	(0.040)	-0.066	(0.058)	-0.175***	(0.066)
7. Windsorize ENTS at 0.10 (top 1.7%)	-0.271***	(0.105)	-0.088	(0.154)	-0.384**	(0.155)
Panel B: Hours Regressions (INTS)						
1. Baseline	0.401**	(0.157)	0.694***	(0.174)	-0.334	(0.310)
2. Smoothed Retirement Incentive Controls	0.398**	(0.162)	0.685***	(0.176)	-0.307	(0.320)
3. Earnings-Based Retirement Definition	0.383**	(0.163)	0.720***	(0.182)	-0.372	(0.307)
4. Retirement Definition Based on Self-reports	0.144	(0.151)	0.410**	(0.162)	-0.450	(0.319)
5. Hours≥10	0.327*	(0.168)	0.691***	(0.181)	-0.602*	(0.353)
6. Hours≥20	0.295**	(0.150)	0.528***	(0.169)	-0.252	(0.294)
7. Windsorize INTS at 0.10 (top 1.2%)	0.222	(0.149)	0.429***	(0.163)	-0.254	(0.305)
Panel C: Earnings Regressions (INTS)						
1. Baseline	0.097	(0.221)	0.006	(0.264)	0.442	(0.429)
2. Smoothed Retirement Incentive Controls	0.145	(0.228)	0.010	(0.277)	0.509	(0.430)
3. Earnings-Based Retirement Definition	0.227	(0.235)	0.144	(0.279)	0.477	(0.463)
4. Retirement Definition Based on Self-reports	-0.104	(0.228)	0.113	(0.279)	-0.309	(0.476)
5. Cut-off $5000	0.016	(0.203)	-0.054	(0.242)	0.293	(0.393)
6. Cut-off $1000	0.138	(0.239)	0.019	(0.288)	0.664	(0.468)
7. Windsorize INTS at 0.10 (top 1.2%)	0.128	(0.257)	-0.081	(0.309)	0.751	(0.499)

Notes: Independent Variable is ENTS in Panel A and INTS in Panels B and C. All dollars are 2003 dollars. Standard errors, clustered by individual, are in parentheses. * indicates p-value<.10; ** indicates p-value<.05; *** indicates p-value<.01. Regression controls are as in Table 5. Probit regressions report marginal effects. Windsorizing ENTS at 0.10 affects 1.7% of observations. Windsorizing INTS at 0.10 affects 1.2% of observations. Except where noted, regressions in Panel B are limited to individuals who work at least 15 hours in a typical week. "Hours≥10" regressions limit the sample to individuals who work at least 10 hours in a typical week, likewise for "Hours≥20" regressions. Except where noted, regressions in Panel C are limited to individuals with at least $2,500 in earnings in the sample year. "Cut-off $5,000" regressions limit the sample to individuals reporting at least $5,000 of earnings in the year of observation, likewise for "cut-off $1,000" regressions. Retirement definitions are explained in Appendix 1.

Sample Selection Criteria

Reason	Persons Dropped	Persons Remaining	Person*Years Dropped	Person*Years Remaining
PANEL A:				
Initial HRS observations		30,207		
Not in original cohort of HRS	17,625	12,582		
Not linked to SSA records	3,113	9,469		
Born before 1920 (different Social Security rules)	45	9,424		
Reported as married in initial wave & spouse is present, but spouse not linked to SSA records	1,087	8,337		49,897
Retired before wave of observation (if before wave 1, all person*years dropped)	1,421	6,916	20,201	29,696
Missing education or race data	9	6,907	38	29,658
Respondent or spouse younger than 40	212	6,695	1,021	28,637
Respondent or spouse reports ever being disabled	581	6,114	1,850	26,787
More than 10 years of non-FICA-covered work. (Possibly subject to Windfall Elimination Provision)	81	6,033	469	26,318
Reports being unmarried, but reports spousal data	57	5,976	402	25,916
Widowed, but no information on deceased spouse	252	5,724	1,245	24,671
Marital status reported as divorced or separated (and either never remarried or remarried after age 60)	725	4,999	3,374	21,297
Respondent is younger than 52	228	4,771	2,660	18,637
Weak labor force attachment prior to entering HRS (2 of past 5 years or previous year earnings < $2500)	800	3,971	4,735	13,902
Resulting Sample for Retirement Regressions		**3,971**		**13,902**
PANEL B:				
Retirement Regressions Sample		3,917		13,902
Respondent's predicted earnings in year of observation is above SS taxable maximum	302	3,669	1,687	12,215
Respondent reports fewer than 15 hours per week in year of observation	517	3,152	1,375	10,840
Resulting Sample for Hours Regressions		**3,152**		**10,840**
PANEL C:				
Retirement Regressions Sample		3,917		13,902
Respondent's predicted earnings in year of observation is above SS taxable maximum	302	3,669	1,687	12,215
Respondent reports less than $2500 in earnings in year of observation	202	3,467	1,153	11,062
Resulting Sample for Earnings Regressions		**3,467**		**11,062**

Appendix Table 2
Sets of Control Variables

Specification	Z_t (1) Determinants Of Labor Supply Not in SS Benefit Calculator (48 vars)	Components of X_{t-1} (2) Last Year's Earnings (2 vars)	(3) Demographics Relevant to SS Benefit Calculator (102 vars)	(4) Spousal Variables Relevant to SS Benefit Calculator (200 vars)	(5) Earnings 2-52 Years Ago (102 vars)	Interactions and higher-order terms of X_{t-1} (6) Aggregates of Yearly Earnings (12 vars)	(7) Higher Order Terms (18 vars)	(8) Interactions of Education and Lifetime Earnings (64 vars)	(9) Dummies for Last Year's Earnings Percentile (99 vars)	(10) Interaction of Previous 5 Years' Earnings (320 vars)	Total Variables (excluding colinear terms)
No Controls											0
Only Basic Demographics		X	Partial								12
Baseline Minus Earnings History	X	X	X	X		Partial					303
Baseline Minus Higher Order Terms	X	X	X	X	X	X					415
Baseline Controls	X	X	X	X	X	X	X				433
Baseline Plus Additional Interactions	X	X	X	X	X	X	X	X	X		581
Baseline Plus Further Additional Interactions	X	X	X	X	X	X	X	X	X	X	836

Notes: Numbers of variables in parentheses in each column include those dropped due to collinearity, but the Total Variables column excludes them. Earnings are real personal SS-eligible earnings in 2003 dollars. Controls for dollar-denominated variables consist of the log of that variable for amounts (in 2003 dollars) of at least $1000 and zero otherwise, and an indicator for the amount being at least $1000. Details on the exact control variables in each column follow.

Column 1: Determinants of Labor Supply Not in SS Benefit Calculator include 2 variables for family assets, 10 dummies for Census region of residence, 32 dummies indicating the longest industry and occupation of the respondent, a dummy for veteran status, a dummy for having been born in the USA, and 2 variables for own and spousal tenure at the longest job.

Column 2: Last Year's Earnings include a dummy for whether indexed Social Security earnings were at least $1000 in the previous year and a variable for the amount of indexed Social Security earnings.

Column 3: Demographics Relevant to the SS Benefit Calculator include 3 dummies for race, 4 dummies for education (no HS degree, HS degree, some college, college graduate), 4 dummies for marital status, 7 dummies for year of observation, 82 age*sex dummies, and 2 for a quadratic in age (which is collinear in all but the "Only Basic Demographics" specification). Race, education, and sex enter the calculator because life expectancy depends on these variables. In the "Only Basic Demographics" specification, only dummies for education, dummies for marital status, dummies for race, and a quadratic in age are included.

Column 4: Spousal Variables Relevant to the SS Benefit Calculator include 2 variables for whether and the number of years the spouse has been retired, 104 variables for earnings for each of the past 52 years, 86 age*sex dummies, 1 variable for the age difference with the respondent, 4 education dummies, and 3 race dummies.

Column 5: Earnings 2-52 Years Ago include 102 variables for earnings for each of the past 52 years except the most recent year.

Column 6: Aggregates of Yearly Earnings consist of 2 variables for Expected SS Wealth, 4 variables for lifetime earnings interacted with sex, 4 variables for spousal lifetime earnings interacted with sex, 2 variables for the number of years worked in total by the respondent and spouse. Lifetime earnings consist of the present discounted value of all personal SS-eligible earnings (using a 3% real discount rate). In the "Baseline Minus Earnings History" regressions, only the Expected SS Wealth variables are included.

Column 7: Higher Order Terms are 2 variables of second and third order terms of Social Security wealth, 8 variables of second and third order terms of lifetime income interacted with sex for both the respondent and spouse, 4 variables of second order terms of current job tenure for the respondent and spouse, and 4 varibles of second and third order terms of total years of labor force experience for the respondent and spouse.

Column 8: Log of lifetime earnings and education dummies are interacted with dummies for each separate marital status by sex and by age category.

Column 9: Last year's earnings percentile are based on personal (SS-eligible) earnings.

Column 10: Interaction of Previous 5 Years of Earnings denotes dummies for each separate marital status (and whether spouse is retired, in the case of married individuals) by sex and by age category (0-62, 63-65, 66-69, and 70+) cell interacted with the 10 variables measuring earnings in the past 5 years.