Preliminary and incomplete

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

Jeffrey B. Liebman
Erzo F.P. Luttmer
David G. Seif

July 11, 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 incentive 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 1.7 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 0.36 and statistically significant, but we do not find a statistically significant earnings elasticity.

---

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

1. Introduction

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 OASDI payroll tax as a pure tax—they fail to recognize that the payment of Social Security taxes will increase their future Social Security benefits. With personal retirement accounts, in contrast, the link between contributions and future income would be clear, and the economic distortions would be reduced. 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 on the effect of Social Security on retirement, Diamond and Gruber (1999) note that most of this literature ignores the effect of the marginal Social Security benefit rate (focusing instead on the 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 to the literature.

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 explictly 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 retirement patterns (Gustman and Steinmeier 1986, 2005a, Rust and Phelan 1997, Laitner and Silverman 2006). 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." Research that uses cross-sectional variation in incentives can deal with this the concern by including determinants of these incentives as control variables. This 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. They will be unbiased if unobserved determinants of labor supply are uncorrelated with these non-linearities. This is more likely in cases where 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 being 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 variation 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, 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, individuals have the choice of claiming 100 percent of own benefits or 50 percent of spousal benefits, creating a discontinuity in marginal benefits around the point where the retired-worker 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 Social Security benefits. Fourth, the kink points in the AIME-PIA conversion schedule[1] 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.

All of these five provisions 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 create non-linearities that will 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 causal effects of the effective Social Security tax.

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 born between 1931 and 1941 as well as their spouses. The original cohort has been

---

[1] The AIME-PIA conversion schedule is explained in section 2.1, where we describe the Social Security benefit rules.

[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 Data file (2008).

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 1.7 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 0.36 and statistically significant, but we do not find a statistically significant earnings elasticity. 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 somewhat more sensitive to changes in specification. Though we lack statistical power to precisely estimate results within subsamples, 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 allow us to clearly 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 accurately determine to what degree individuals perceive this tax-benefit link. They are both consistent 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. 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 are 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 at older ages are not indexed. A worker's average indexed monthly earnings (AIME) are calculated by summing the worker's indexed earnings in 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 benefit age and claims benefits as a retired worker. The PIA is indexed for inflation. Workers may claim benefits as early as age 62, with a reduction in benefits of about 6 2/3 percent per year. Workers who delay claiming beyond the full benefit age receive increased benefits from the delayed retirement credit. In married couples, the lower-earnings spouse receives the maximum of 50 percent of the benefit of the higher-earnings spouse and the spouse's own retirement benefit. Surviving spouses in married couples receive a benefit equal to the maximum of the retired worker benefits of the two 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 the following five provisions that generate the variation we exploit in our empirical analysis.[3]

---

[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 are planning to claim spousal or widow benefits. (3) Remarriage –

First, we exploit the fact that Social Security benefits depend on only the 35 highest years of indexed earnings (the "maximum-35-year rule"). Thus, after 35 years of earnings, an additional year of earnings will increase benefits only inasmuch as the additional year of earnings exceeds a year of lower earnings. This implies that the effective marginal Social Security benefits from an additional year of work are lower for an individual the higher are the earnings in the year replaced. If the earnings in the additional year of work exceed the lowest earnings among the 35 highest years of earnings (as is typically the case), then the *average* returns from working the additional year are reduced relative to someone who was replacing a year with zero earnings. 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. The 35-year rule will reduce the marginal returns to working if the additional hours take place in a year that is currently not among the 35 highest years of earnings, or when there is some chance, given uncertainty about future earnings, that the current year will no longer be among the 35 highest years of earnings at the point when the person's Social Security benefits are calculated.

Second, the rules on spousal benefits create variation in the effective Social Security tax rate. Again, this variation consists of non-linearities that converge to discontinuities as uncertainty about future own and spousal labor supply gets resolved. Individuals have a choice between claiming benefits based on 100 percent of their own Primary Insurance Amount (PIA) or based on 50 percent of their spouse's PIA. 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

---

individuals lose eligibility for spousal benefits based on an ex-spouse upon remarriage, thus creating jumps in marginal Social Security benefits upon remarriage for the subgroup of individuals who would have claimed benefits based on a previous spouse's earnings. (4) The Windfall Elimination Provision – this provision places workers who receive a 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 they have earned from SS-ineligible government work if they are taking any Social Security, thus effectively forcing many workers not to take Social Security benefits. (6) The "Special Minimum PIA" – this creates variation in effective marginal Social Security benefit rates for workers with a similar lifetime earnings but with different earnings histories. (7) Children's benefits – 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.

labor earnings and Social Security benefits. A second 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. This will discontinuously increase the tax-benefit linkage on the margin by about 50 percent.[4]

Third, there is variation due to rules on benefits for widows and widowers. An individual with a living spouse has the choice of claiming her own retired-worker benefit or 50 percent of her spouse's benefit, while someone with a deceased spouse can choose between claiming 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 one's own benefits exceed 50 percent of their spouse's benefits and the spouse is alive in that year, plus the probability that one's 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 drops 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 in the value of the PIA ratio at the point of benefit claiming and turn the discontinuity in the returns to work at the earnings level where the PIA ratio becomes unity into a non-linearity.

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, 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 end up on, the returns to work will be a weighted average of the

---

[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 full retirement age. In other cases, differences in life-expectancy, and early retirement credits or delayed retirement benefits can cause this increase to be somewhat larger or smaller than 50 percent.

[5] As with spousal benefits, early retirement credits or delayed retirement benefits may result in slightly different values.

returns to work at each of the segments weighted by the probabilities of ending up on each of the segments. In other words, for those with uncertainty about future earnings, the kinks in the AIME-PIA schedule will generate 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 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 the interaction effects of 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 (see, e.g., 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 this person: $SSW_{it}(\boldsymbol{X}_{i,t\text{-}1}, \boldsymbol{X}^{+}_{i,t\text{-}1})$, where $(\boldsymbol{X}_{i,t\text{-}1}, \boldsymbol{X}^{+}_{i,t\text{-}1})$ together is the vector of individual characteristics (including own and spousal earnings) that determine Social Security benefit payments. This vector consists of a component, $\boldsymbol{X}_{i,t\text{-}1}$, that is known at time *t*-1, and a component, $\boldsymbol{X}^{+}_{i,t\text{-}1}$, that is not yet known at that time (except under perfect foresight). The person would face an effective Social Security tax of:

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

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}:

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

where $\boldsymbol{Z}_{it}$ is a vector of other explanatory variables for labor supply, while α, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ are parameters to be estimated, and ε is an error term. The functional form of the net-of-tax share, $1 - \tau_{it}^{effective}$, 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:

$$(3)\ h_{it} = \alpha\left(1 - \mathrm{E}[\tau_{it}^{effective} \mid \boldsymbol{X}_{i,t\text{-}1}]\right) + f(\boldsymbol{X}_{i,t\text{-}1}, \boldsymbol{\beta}) + \boldsymbol{Z}_{it}\boldsymbol{\gamma} + \varepsilon_{it}\ ,$$

Due to the expectation operator, E[], discontinuities in the effective marginal tax rate turn into non-linearities.[6] These non-linearities would be fully absorbed by *f*(.) if we allow *f*(.) to be an arbitrarily flexible function of $\boldsymbol{X}_{i,t\text{-}1}$ 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 ($\boldsymbol{X}_{i,t\text{-}1}$) and unobserved

---

[6] Note that some forms of uncertainty preserve discontinuities. For example, the expectations operator integrates over all possible ages of own death, but variation in the age of death after claiming benefits does not smooth out discontinuities.

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 those provisions that create the discontinuities. We refer to the Social Security rules stripped of the provisions that create discontinuities as the "smoothened" Social Security benefit rules. In particular, we (i) eliminate the 35-year rule by letting the smoothened 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 that people in our data set that have the same sex, own work status, marital status and spousal work status, receive on average, (iii) replace the kinked AIME-PIA schedule by the best-fitting quadratic schedule, and (iv) eliminate the vesting rule. The resulting "smoothened" Social Security rules closely resemble the actual rules, except that they no longer contain discontinuities.

We then use these smoothened rules to calculate a smoothened expected effective Social Security tax rate ($\tau_{it}^{Smoothened}$) 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:

$$(4)\ h_{it} = \alpha\left(1 - \mathrm{E}[\tau_{it}^{Smoothened} \mid \boldsymbol{X}_{i,t\text{-}1}]\right) + f(\boldsymbol{X}_{i,t\text{-}1}, \boldsymbol{\beta}) + \boldsymbol{Z}_{it}\boldsymbol{\gamma} + \varepsilon_{it}\ ,$$

In these regressions, the effect of the smoothened 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 the benefit payments are less far in the future for older individuals). While 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 and then use this 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 4) is driven by the variation in effective tax rates that is due to 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.

3. Data and Empirical Implementation

3.1 Data

We perform our estimation using observations from the original cohort of the Health and Retirement Study (HRS), a longitudinal survey that can be linked to Social Security 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 13,367 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.[7] 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.

The HRS also indicates self-reported retirement status as well as the year and month that each individual retired (if the individual reports being retired), allowing us to construct a measure of retirement status at each survey date. In some cases, individuals report being retired but nevertheless report substantial labor earnings after their retirement date. In those cases, we infer retirement based on the actual labor earnings record. We assume the retirement year after earnings fell permanently below $2,500.[8] The HRS survey data also contains the two other dependent variables for our regressions: earnings and hours worked per week. The first of these is directly 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 correspond to the year prior to the survey year.

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 to administrative Social Security records (about one-third of potential observations). We also exclude individuals who were already retired as of the initial wave of the HRS or who had very weak past labor force attachment (about 14 percent of potential observations). In addition, we exclude widowed, separated, and divorced individuals in cases in which we have insufficient information about their former spouses to calculate benefits (about 9 percent of potential observations). Other sample restrictions result in much smaller numbers of dropped observations. This leaves us with

---

[7] Social Security benefits do not depend on earnings earned in years prior to 1951.

[8] 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.

a sample of 3,971 individuals (2,269 men and 1,702 women) out of the 13,367 individuals in the original HRS cohort.

We limit our sample to person-year observations of 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, a hazard rate that does not significantly vary by sex. Conditional on working, the average male worker works almost 42 hours per week, while the average female worker works 35 hours per week. Mean Social Security wealth is around $270,000. 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 are people who are married. 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 the two-year retirement hazard rate 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 from ages 60 to 62 and then remains relatively constant thereafter.[9]

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)[10] 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

---

[9] We do not show retirement hazards beyond the age of 70 because we have fewer than 100 observations in each age-gender cell off of which to estimate hazard rates for ages 71 or greater.

[10] We exclude the disability insurance component of OASDI because DI benefits are not incorporated into our model.

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, the early retirement penalty, the delayed retirement credit, and the vesting rule based on quarters of earnings.[11] We assume that individuals claim the higher of the benefits they are entitled to on their own record and the benefits they are entitled to on their living or deceased spouse's record. We assume this decision is made (or updated) when (i) 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.[12] Further details of this calculation are spelled out in Appendix 1.

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 started claiming benefits, and (x) year in which the spouse started claiming benefits. Year of birth and earnings history are known,

---

[11] 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. These provisions would apply to relatively few observations. In order to include them, we would need information 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.

[12] Our practice 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.

but the remaining six variables are generally stochastic.[13] Thus, future Social Security benefits are an expectation with respect to six 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 benefit 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 retirement occurs no later than at age 80 and that death occurs no later than at age 100. 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. 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 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.

---

[13] 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.

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:

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

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

$$(6) \qquad ENTS_{it} = \ln\Big(1 - 0.106 / 1.053 + \big(ESSW_{it}(\text{retire in } t+1) - ESSW_{it}(\text{retire in } t)\big) / \hat{y}_t\Big).$$

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 with 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

---

[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.

[15] We recognize, but do not model, the option value in the decision not to retire, as highlighted by Stock and Wise (1990).

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 take-up decision) and the level of future benefits

4. Results

4.1 Effective Social Security Net-of-Tax Shares

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 their own right, because they inform how benefit rules could be restructured to reduce the size and variation in distortions. Indeed, this is the focus of a number of papers in the literature, see for example Feldstein and Samwick (1992), Goda (2007), Sabelhaus (2007), and Goda, Shoven, and Slavov (2009). Second, we want to document the variation in the incentives. If our estimated incentives correspond to those of previous papers and 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. Future benefits are discounted to the present using a 3% real discount rate. The distribution of Social Security Wealth is skewed slightly to the left with a mode

---

[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.

[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.

around \$290,000. 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. The second and third columns of Table 2 shows 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 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. Thus, on average, the effective Social Security tax is 6.9 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.

The fourth and fifth columns of Table 2 shows the mean and standard deviation of INTS by demographic subgroup. The variation in INTS by subgroup is generally in line with previous studies that calculate the intensive-margin incentives from Social Security (Feldstein and Samwick, 1992; Armour and Pitts, 2004; Cushing, 2005; Goda, 2007). Work incentives are lower for women than for men because women are more likely to claim off of their spouse's record. Among men, work incentives are stronger for those with shorter work histories, and lower life-time 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' record.

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 have no tax-benefit linkage, for example because their current predicted indexed earnings are not among their 35 highest annual indexed earnings. On the other extreme, 1.7 percent of observations have very high tax-benefit linkages because one additional year of earnings will give them sufficient quarters of earnings such that they will qualify for receiving Social Security benefits. 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 3,329% percent on earnings, or an ENTS of 3.50. While these incentives are most likely real, they are clearly outliers.[18] This produces a risk that the regressions will be driven by the handful of observation that qualify for Social Security because they reach the required number of quarters of earnings. To avoid this, we topcode ENTS at 0.5, which is slightly above the highest value of ENTS achieved by someone with more than 40 quarters of earnings. We will

---

[18] 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.

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 lies under 0.001.

Columns 6 and 7 of Table 2 shows the mean and standard deviation of ENTS by demographic subgroup. The variation in ENTS by subgroup is generally in line with previous studies that calculate the extensive-margin incentives from Social Security (Butricia et al. 2006; Goda, Shoven, and Slavov, 2006; Goda 2007; Sabelhaus 2007). Even though the INTS and the ENTS are often quite different for individuals in certain years, the 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 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 the intensive-margin incentives (INTS). 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 percentage 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 20 percent 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 70 percent and 170 percent, 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 220 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.[19] 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 70 percent, where

---

[19] 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 the minimum. 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 be zero.

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 have converged to true discontinuities, as the woman's date of retirement is no longer 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 $f(\boldsymbol{X}_{i,t-1}, \gamma)$ 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 *smoothened* ENTS and INTS on increasingly flexible sets of control variables. Recall that the smoothened 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 smoothened incentives.

Our baseline control function $f(\boldsymbol{X}_{i,t-1}, \beta)$ 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.[20] In addition, the linear combination contains a full set of dummies for own gender×age, for spousal gender×age, for the individual's marital status, 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

---

[20] 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 real dollars. We do so for all variables in this paper that are logs of dollar amounts.

black, non-Hispanic white, 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 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 quadratic polynomial in the number of years in which own Social Security earnings exceeded $1000. In total, $f(\boldsymbol{X}_{i,t-1}, \boldsymbol{\beta})$ is a linear combination of 382 terms. The baseline vector of additional controls, $\boldsymbol{Z}_{it}$ consists of the log of household assets, a quadratic 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, $\boldsymbol{Z}_{it}$ 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 smoothened 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 smoothened ENTS is 4.55 percentage points, while the standard deviation for the smoothened 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 94 percent of the variation in the smoothened 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 smoothened incentives (Appendix Table A1 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 75%. Rows 6 adds an additional 141 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 255 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 smoothened ones because of the additional variation due to the discontinuities-in-the-limit. Conversely, the R^2 is lower in all specifications. The baseline set of controls is only able to explain about 57 percent of the variation in the extensive-margin incentives and about 67 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 94% to almost 98%, 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 smoothened incentive variables insignificant. In this table, we show regressions of a measure of labor supply (retirement, hours, or earnings) on the smoothened 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 both correlated with these discontinuities-in-the-limit and with labor supply. We therefore consider 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 430 baseline control variables. The first column shows the estimates for the whole sample, while the next two columns separate the estimates by gender. We find that, in the whole sample, 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 1.7 percentage points on a base of 15.1, or about 12 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 signficant 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 extentive-margin incentives can only be rejected at a marginal level of significance (p-value: 0.07).

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 remain in the labor market in order to be included in the regression.

This sample restriction is implemented as 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 0.36 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 0.19. 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, -0.33 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 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. 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 only 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, and Kopczuk 2005). Given this range and the confidence intervals on our estimates, our estimates are consistent with a full perception of the tax-benefit link and with only perceiving a minute fraction 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 specifications with less extensive specifications of the control function, and as a result these estimates may be driven by variation in the incentives that is not due to the five provisions in 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 for men in row 2), confirming that some of variation that is unrelated to the discontinuities-in-the-limit may indeed be 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 controls variables roughly doubles as we go 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 perhaps 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 we observe the sharpest discontinuities 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 loses statistical significance in rows 7 and 8, though the hours elasticity for men retains it statistical significance. The finding that the hours elasticity is 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, respectively, the retirement regression, the hours regression, and the earnings regression. The first row of each panel reproduces the baseline regression. The second row show that results are very robust to including the smoothened 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, namely ones 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 quite sensitive to the choice of retirement definition. While a retirement definition that combines both sources of information is arguably more precise (which is why we use this definition in the first place), the sensitivity of the results to the retirement definition is nevertheless an important caveat to the 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 significance 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 about a third causing this estimate to lose statistical significance, though the estimate in the subsample of men is still significant at the 5-percent level.

5. Discussion

Estimating how the effective marginal Social Security tax affects labor supply is challenging because this 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. Since we believe these discontinuities-in-the-limit yield credible variation in labor supply incentives, our methodology relies only on this variation to identify the labor supply response to the effective Social Security tax rate.

Our estimates clearly 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 imply that the potential efficiency gains from increasing the transparency of the link between Social Security benefits and taxes are 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, but clearly this is a fruitful area for future research. 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

Armour, Brian S., and M. Melinda Pitts. 2004. “Incorporating Insurance Rate Estimates and Differential Mortality into the Net Marginal Social Security Tax Rate Calculation,” *Public Finance Review*, 32(6), pp. 588-609.

Auerbach and 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. 243-252.

Burkhauser, Richard V., and John A. Turner. 1985. “Is the Social Security Payroll Tax a Tax?” *Public Finance Quarterly*, 13(3), pp. 253-267.

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), 253-266.

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

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. 245–262.

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.

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., and Thomas L. Steinmeier. 1986. "A Structural Retirement Model," *Econometrica*, 54(3), pp. 555-584.

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. 441-463.

Gustman, Alan L., and Thomas L. Steinmeier. 2005b. " Imperfect Knowledge of Social Security and Pensions," *Industrial Relations*, 44(2), pp. 373-397.

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. 201-209.

Kopczuk, Wojciech. 2005 "Tax Bases, Tax Rates and the Elasticity of Reported Income," *Journal of Public Economics*, 89(11-12), pp. 2093-2119.

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. 412-437.

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

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. 2008. 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. 207-236.

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

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. 135-155.

Appendix 1: Calculation of Expected Social Security Wealth

We define family Social Security Wealth for an individual and his or her spouse combined. Conceptually, it is the expected present discounted value of all future payments from the Social Security Administration to the individual and his or her spouse.

Let $B(i)$ denote the date of birth of individual i, which fully defines the year the individual reaches the Full Retirement Age (*FRA*) and the Early Retirement Age (*ERA*). Let $\mathbf{y}_{it} = \{\ldots, y_{i,t-2}, y_{i,t-1}\}$ denote the individual's earnings history up to year t (but excluding year t), and let $\mathbf{y}_{it}^{+} = \{y_{it}, y_{i,t+1}, y_{i,t+2}, \ldots\}$ denote earnings in year t and in future years. Unless otherwise noted, all monetary variables are expressed in real 2003 dollars. Let $PIA_{i,t+n,B}(\mathbf{y}_{it}, \mathbf{y}_{it}^{+})$ be the Primary Insurance Amount (*PIA*)[21] of individual i in year $t+n$ as a function of the individual's earnings profile and date of birth. The spousal analogues of the above are $\mathbf{y}_{it}^{S}, \mathbf{y}_{it}^{S+}$, and SPIA.

The PIA is determined by first calculating the AIME in the same year and converting this AIME to the PIA using the AIME-PIA conversion schedule (which does not vary by year for the same individual). The AIME is calculated by taking the average of the 35 highest years of monthly indexed AIME-eligible earnings for individuals born in 1928 or later, and of the $35-1928+B$ highest years (where B is, in this case, *year* of birth only) for individuals born between 1920 and 1928. In addition, an individual's AIME is set to zero if he has too few quarters of earnings, as defined by the Social Security Administration.[22]

Let $D(i)$ denote the year in which individual i dies, let $SB(i)$ be individual i's spouse's date of birth[23], let $SD(i)$ be individual i's spouse's date of death, and let $DPIA_{i,B,D,SB}(\mathbf{y}_{it}, \mathbf{y}_{it}^{+})$ be the Death Primary Insurance Amount (*DPIA*), the PIA bequeathed to one's spouse upon death, of individual i as a function of the individual's earnings profile, date of birth, year of death, and spouse's date of birth. When calculating the

---

[21] *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.

[22] Rules regarding the effect of disability on one's PIA are ignored.

[23] In the event of individuals who never married, all spousal date values are $\varnothing$, and all probabilities may be set to any values so long as they satisfy all equations listed in this section. These $\varnothing$ values are in addition to any other conditions listed whereby these values are $\varnothing$ or zero.

DPIA, all inputs of $\mathbf{y}_{it}$ and $\mathbf{y}_{it}^{+}$ after death must be zero. Note that whenever an individual dies at the age of 62 or greater, the PIA and the DPIA are the same (given $\mathbf{y}_{it}$ and $\mathbf{y}_{it}^{+}$). For individuals who die before turning 62, the DPIA is different from the PIA due to the number of years included and the year the wage indexing of earnings stops, as per Social Security Administration rules. In addition, given the date of an individual's death and all other variables, DPIA does not vary over time. The spousal analogue of DPIA is SDPIA.

Let p_{it} denote the probability that individual i is alive in year t[24], based on the age- and gender-specific cohort life tables used by the Social Security Administration[25], and let d_{it} be the probability the same individual dies in period t. These probabilities are related such that for any $k \geq 0$:

$$p_{i,t+k} = 1 - \sum_{j=0}^{k} d_{i,t+j}$$

The spousal analogues of these probabilities are p_{it}^{S} and d_{it}^{S}.

In year t, future earnings (including those in year t) are not yet known. We therefore treat them as stochastic variables. In particular, we calculate the age- and gender-specific probabilities of future labor force participation based on the age- and gender-specific retirement and death hazard rates. We calculate expected future earnings conditional on being in the labor force by applying the age- and gender specific earnings growth to each year's earnings. Let r_{itj} denote the probability that individual i's earnings ended in year t[26] as a result of retirement conditional on death in year j, and let u_{itj} denote the probability that individual i's earnings ended in year t as a result of death conditional on death in year j. This condition implies that $r_{itj} = 0, \forall t > j$. However, for all values of $t \leq j$, r_{itj} is uncorrelated with j. The probability from a situation whereby one both retires and dies in the same year accrues to r_{it}, as retirement clearly happened first, and

---

[24] Due to our discretization of time, we define being alive "in" t as entering *and* exiting the period alive. Being alive in t is a requirement in order to collect Social Security benefits or work in t, but one need only *enter* the period alive (i.e., have been alive in t-1) in order to retire, die, or take up Social Security benefits in t. In the event that one both takes up benefits and dies in t, no benefits are actually collected. In addition, an individual whose spouse dies in t is eligible for widow(er) benefits in t instead of spousal benefits (assuming, of course, that the individual himself is alive in t).

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

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

death was thus not the cause of the end of the earnings. The variable u_{itj} therefore only captures the probability that one's earnings end in t and one dies in t, too, having never retired. In effect, u_{itj} captures the sum of the retirement probabilities that *would have existed* into the future if death had not occurred in j. It is clear that $u_{itj} = 0, \forall t \neq j$, as one cannot have an earnings history that ends after death, and an earnings history that ends before death is due to retirement. These probabilities, r_{itj} and u_{itj}, are defined such that, along with d_{it}:

$$1 = \left(\sum_{k=0}^{j} r_{i,t+k,t+j} \right) + u_{i,t+j,t+j}, \forall j \geq 0$$

and

$$1 = \sum_{j=0}^{\infty} \left[d_{i,t+j} \left(\left(\sum_{k=0}^{j} r_{i,t+k,t+j} \right) + u_{i,t+j,t+j} \right) \right]$$

Additionally:

$$r_{itj} = r_{itk}, \forall t < j,k$$

The spousal analogues of the above are r^{S}_{itj} and u^{S}_{itj}.

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(er) 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 retirement age (if the individual has retired) and, when taking spousal or widow(er) benefits, his or her spouse's date of birth, retirement age (if the spouse ever retired before dying), and age of death. Let $R(i)$ be the year of retirement for individual i, and let $SR(i)$ be the year of retirement for individual i's spouse. The values for R and SR are $\varnothing$ in cases where retirement never takes place before death.

Let the age of take-up of Social Security benefits for individual i be $T_{B,R,D}$. This is equal to the greater of 62 and R, but is capped at 70. If take-up never occurs (due either to death before age 62 or both death before age 70 and non-retirement), T is $\varnothing$. Let $ST_{SB,SR,SD}$ be the same for individual i's spouse.

Let $A_{t,B,D,R}$ be the benefit adjustment factor in year t for an individual taking benefits based on his or her own earnings, given a birth date of B, a retirement date of R, and a death date of D. This factor is zero in years before the early retirement age is reached and after death. Note that these four inputs fully define the age of take-up and whether take-up ever occurs. A is equal to zero if take-up has not yet occurred or if t is greater than D. The actual benefits received are equal to the adjustment factor times the PIA.

Let $SA_{t,B,D,R,SB,SD}$ be the adjustment factor in year t for an individual taking spousal benefits (benefits based on his or her *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, SB, and of *i*'s spouse's date of death, SD. SA is zero if any of the following are true in *t*: *i*'s spouse has not reached the ERA, *i* has not taken up benefits, *i* has died, or *i*'s spouse has died. The actual benefits received are equal to the adjustment factor times half of the spouse's PIA.

Let $WA_{t,B,D,R,SB,SD,SR}$ be the adjustment factor in year t for someone taking widow benefits. This is a function of everything of which SA is a function, as well as of *i*'s spouse's retirement date. 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 $t+n, n \geq 0$ is found by taking the expectation over all possible combinations of the own and spousal labor exit date and own and spousal date of death of the maximum benefits in that year. It is also necessary to differentiate between the two possible reasons for exiting the labor force, voluntary retirement and death. Given two possibilities per person and two person couples, there are four possible "retirement combinations," each of which contributes to the expected Social Security benefit in a given year. Conditional on earnings histories, 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 due to both individuals voluntarily retiring is $E\left[SSB_{i,t+n}^{R,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$, which is:

$$E\left[SSB_{i,t+n}^{R,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=j}^{\infty}\sum_{m=k}^{\infty} d_{i,t+l} d_{i,t+m}^{S} r_{i,t+j,t+l} r_{i,t+k,t+m}^{S} 1\left(T_{\substack{B_i, D=t+l, \\ R=t+j}} \geq t+n\right) \cdot$$

$$\Bigg[1(m > n)\max\left\{A_{\substack{t+n,B_i,\\ D=t+l,R=t+j}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=t+j}}, SA_{\substack{t+n,B_i,D=t+l,\\ R=t+j,SB_i,SD=t+m}} \cdot \frac{1}{2} ESP_{\substack{i,t,n,SD=t+m,\\ SR=t+k}}\right\} +$$

$$1(m \leq n)\max\left\{A_{\substack{t+n,B_i,\\ D=t+l,R=t+j}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=t+j}}, WA_{\substack{t+n,B_i,D=t+l,R=t+j,\\ SB_i,SD=t+m,SR=t+k}} \cdot ESDP_{\substack{i,t,SD=t+m,\\ SR=t+k}}\right\}\Bigg]$$

In the above equation, 1() is the indicator function (which equals one if the expression between parentheses is true and zero otherwise), and:

$$EP_{\substack{i,t,n,D=t+l,\\ R=t+j}} = PIA_{B_i,t+n}\left(\mathbf{y}_{it}, \mathbf{y}_{it}^{+} |_{\substack{D=t+l,\\ R=t+j}}\right)$$

$$ESP_{\substack{i,t,n,SD=t+m,\\ SR=t+k}} = SPIA_{B_i,t+n}\left(\mathbf{y}_{it}^{S}, \mathbf{y}_{it}^{S+} |_{\substack{SD=t+m,\\ SR=t+k}}\right)$$

$$ESDP_{\substack{i,t,SD=t+m,\\ SR=t+k}} = SDPIA_{B_i}\left(\mathbf{y}_{it}^{S}, \mathbf{y}_{it}^{S+} |_{\substack{SD=t+m,\\ SR=t+k}}\right)$$

The term $EP_{\substack{i,t,n,D=t+l,\\ R=t+j}}$ is thus the expectation in year *t* of individual *i*'s PIA in *t+n* conditional on individual *i* retiring in year *t+j* and dying in year *t+l*. The first line of the $E\left[SSB_{i,t+n}^{R,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$ equation sums over all possible retirement periods and death periods for both the individual and the spouse, and the indicator function sets benefits to zero if the individual has not yet taken them up. The second line is operative if and only if the individual's spouse is still alive, in which case the individual takes the maximum of his or her own benefits and the spousal benefits to which he or she is entitled. The third line is operative if and only if the individual's spouse is dead, in which case the individual takes the maximum of his or her own benefits and his or her widow(er) benefits. In practice, $E\left[SSB_{i,t+n}^{R,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$ captures the vast majority of expected Social Security benefits, as voluntary retirement is far more common than labor force exit due solely to death. Note that the summations over *l* and *m* could start at zero instead of at *j* and *k*, respectively, without affecting the total, as all terms that would be included by modifying the summations as such would be equal to zero. In addition, single, never married individuals are equivalent to widowed individuals whose spousal benefits and whose widow(er) benefits are always equal to zero whose spouse's death and retirement probabilities are anything that satisfy the conditions laid out earlier.

The portion of benefits in year *t+n* due to voluntary retirement on the individual's part but labor force exit due to death for the individual's spouse is $E\left[SSB_{i,t+n}^{R,SU} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$. Continuing to allow *j*, *l*, and *m* to index year of own retirement, year of own death, and year of spousal death, respectively, and omitting *k*, which previously indexed year of spousal retirement, this is:

$$E\left[SSB_{i,t+n}^{R,SU} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] = \sum_{j=0}^{\infty}\sum_{l=j}^{\infty}\sum_{m=0}^{\infty} d_{i,t+l}\, d_{i,t+m}^{S}\, r_{i,t+j,t+l}\, u_{i,t+m,t+m}^{S}\, 1\left(T_{\substack{B_i,D=t+l,\\ R=t+j}} \geq t+n\right)\cdot$$

$$\left[1(m>n)\max\left\{A_{\substack{t+n,B_i,\\ D=t+l,R=t+j}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=t+j}}, SA_{\substack{t+n,B_i,D=t+l,\\ R=t+j,SB_i,SD=t+m}} \cdot \frac{1}{2} ESP_{\substack{i,t,n,SD=t+m,\\ SR=\varnothing}}\right\} + \right.$$

$$\left. 1(m \leq n)\max\left\{A_{\substack{t+n,B_i,\\ D=t+l,R=t+j}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=t+j}}, WA_{\substack{t+n,B_i,D=t+l,R=t+j,\\ SB_i,SD=t+m,SR=\varnothing}} \cdot ESDP_{\substack{i,t,SD=t+m,\\ SR=\varnothing}}\right\}\right]$$

The portion of benefits in *t+n* due to voluntary spousal retirement but labor force exit due to death on the individual's part is $E\left[SSB_{i,t+n}^{U,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$. Calculating this omits *j*, year of own retirement, from the initial four-summation equation, which yields:

$$E\left[SSB_{i,t+n}^{U,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=k}^{\infty} d_{i,t+l}\, d_{i,t+m}^{S}\, u_{i,t+l,t+l}\, r_{i,t+k,t+m}^{S}\, 1\left(T_{\substack{B_i,D=t+l,\\ R=\varnothing}} \geq t+n\right)\cdot$$

$$\left[1(m>n)\max\left\{A_{\substack{t+n,B_i,\\ D=t+l,R=\varnothing}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=\varnothing}}, SA_{\substack{t+n,B_i,D=t+l,\\ R=\varnothing,SB_i,SD=t+m}} \cdot \frac{1}{2} ESP_{\substack{i,t,n,SD=t+m,\\ SR=t+k}}\right\} + \right.$$

$$\left. 1(m \leq n)\max\left\{A_{\substack{t+n,B_i,\\ D=t+l,R=\varnothing}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=\varnothing}}, WA_{\substack{t+n,B_i,D=t+l,R=\varnothing,\\ SB_i,SD=t+m,SR=t+k}} \cdot ESDP_{\substack{i,t,SD=t+m,\\ SR=t+k}}\right\}\right]$$

Finally, there is the portion of benefits in *t+n* due to both members of a married couple leaving the labor force due to death, which is $E\left[SSB_{i,t+n}^{U,SU} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$. Omitting both *j* and *k* summations, this is:

$$E\left[SSB_{i,t+n}^{U,SU} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty} d_{i,t+l}\, d_{i,t+m}^{S}\, u_{i,t+l,t+l}\, u_{i,t+m,t+m}^{S}\, 1\left(T_{\substack{B_i,D=t+l,\\ R=\varnothing}} \geq t+n\right)\cdot$$

$$\left[1(m>n)\max\left\{A_{\substack{t+n,B_i,\\ D=t+l,R=\varnothing}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=\varnothing}}, SA_{\substack{t+n,B_i,D=t+l,\\ R=\varnothing,SB_i,SD=t+m}} \cdot \frac{1}{2} ESP_{\substack{i,t,n,SD=t+m,\\ SR=\varnothing}}\right\} + \right.$$

$$1(m \le n)\max\left\{ A_{\substack{t+n,B_i,\\ D=t+l,R=\varnothing}} \cdot EP_{\substack{i,t,n,D=t+l,\\ R=\varnothing}}, WA_{\substack{t+n,B_i,D=t+l,R=\varnothing,\\ SB_i,SD=t+m,SR=\varnothing}} \cdot ESDP_{\substack{i,t,SD=t+m,\\ SR=\varnothing}} \right\}\Bigg]$$

Taking all four cases, the total expected Social Security benefit for person *i* in *t+n* is $E\left[SSB_{i,t+n} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$, which is:

$$E\left[SSB_{i,t+n} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] = E\left[SSB_{i,t+n}^{R,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] + E\left[SSB_{i,t+n}^{R,SU} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] +$$

$$E\left[SSB_{i,t+n}^{U,SR} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right] + E\left[SSB_{i,t+n}^{U,SU} \mid \mathbf{y}_{it}, \mathbf{y}_{it}^{S}\right]$$

The expected Social Security benefit in *t+n* for *i*'s spouse, $E\left[SSSB_{i,t+n} \mid \mathbf{y}_{it}^{S}, \mathbf{y}_{it}\right]$, 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 an individual's and an individual's spouse's Social Security benefits. This is:

$$SSW_{it}(\mathbf{y}_t, \mathbf{y}_t^s) = \sum_{n=1}^{\infty}\Big(\big(E[SSB_{i,t+n} \mid \mathbf{y}_t, \mathbf{y}_t^s] + E[SSSB_{i,t+k} \mid \mathbf{y}_t^s, \mathbf{y}_t]\big)(1+\rho)^{-n}\Big),$$

where ρ denotes the real discount rate and all values are real.

One aspect of Social Security that is not included in our calculations is the Special Minimum PIA. Out of our sample of over 10,000 individuals, we found that this would be relevant for fewer than 10.

In addition, as we do with the age of benefit take-up, we define a heuristic for choosing between the each of the two possible benefits (individual versus spousal if married, individual versus widow(er) if widowed) conditional on take-up having occurred. This heuristic requires that an individual never change which benefits he or she is taking except when take-up initially occurs, when a living spouse reaches the ERA, or when a spouse dies. Thus, there are never more than three decisions to be made (and only two involve more than one choice) after taking up benefits. An individual who faces all three choices is one who takes up benefits while his or her spouse is still under the ERA (and whose only "choice" therefore is to take his or her own benefits), then chooses, upon his or her spouse's attainment of the ERA, whether to continue with his or her own benefit or whether to switch to spousal benefits, and chooses once more, upon his or her spouse's death, whether to take his or her own benefit or whether to switch to widow(er) benefits.[27]

This heuristic, while capturing the timing of the most important decisions one makes after take-up, does not allow individuals to "game" the Social Security system, which could potentially yield even higher expected benefits. For example, it may be optimal in expectation for a low-earning spouse to take individual benefits upon reaching the ERA and then switch to spousal benefits upon reaching the FRA, which would allow the individual to avoid a permanent reduction of benefits due to the pre-FRA take-up of benefits. However, including all possible "switches" yields more complexity than can be modeled.

---

[27] In practice, a decision by someone electing to take spousal benefits while his or her spouse is still alive to switch to his or her own benefit (not widow(er) benefits) upon spousal death is virtually never optimal based on this heuristic, as widow(er) benefits are considerably more generous than spousal benefits, and the individual already revealed a preference for spousal benefits over his or her own benefits. Therefore, we assume that any individual taking spousal benefits switches to widow(er) benefits upon spousal death. Without this assumption, the benefit adjustment factor for individual benefits would be a function of more than just the current period and the individual's date of birth, death, and retirement.

Figures

Figure 1
Annualized Two-Year Retirement Hazard

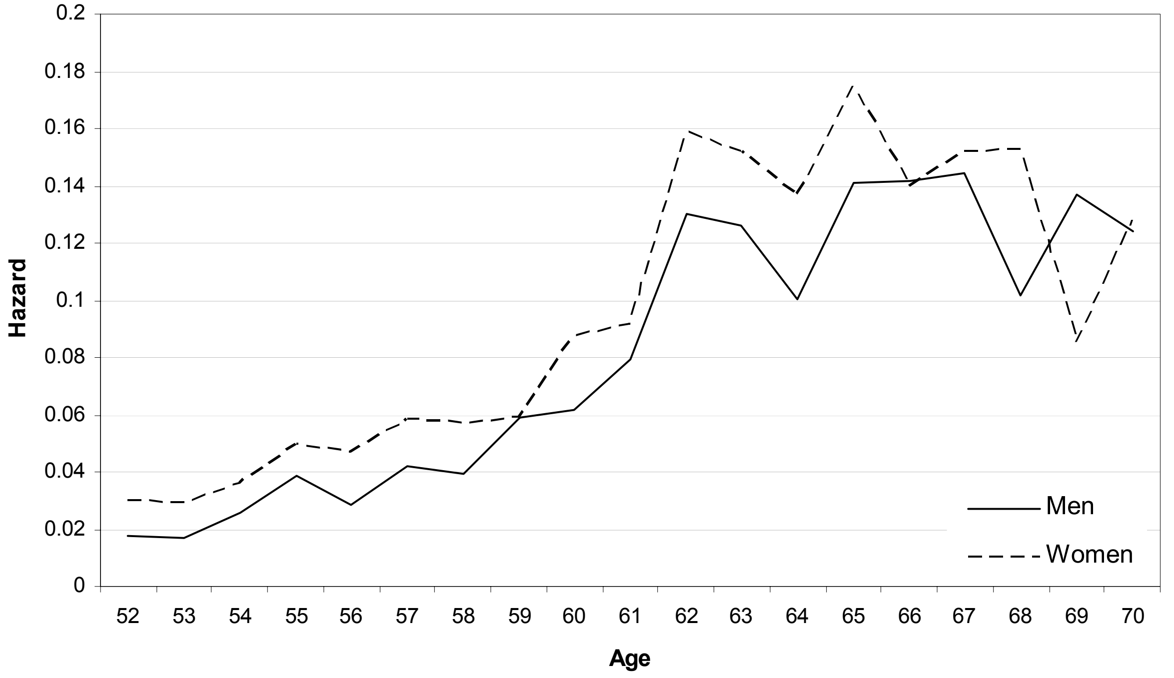


Figure 2
Expected Family Social Security Wealth

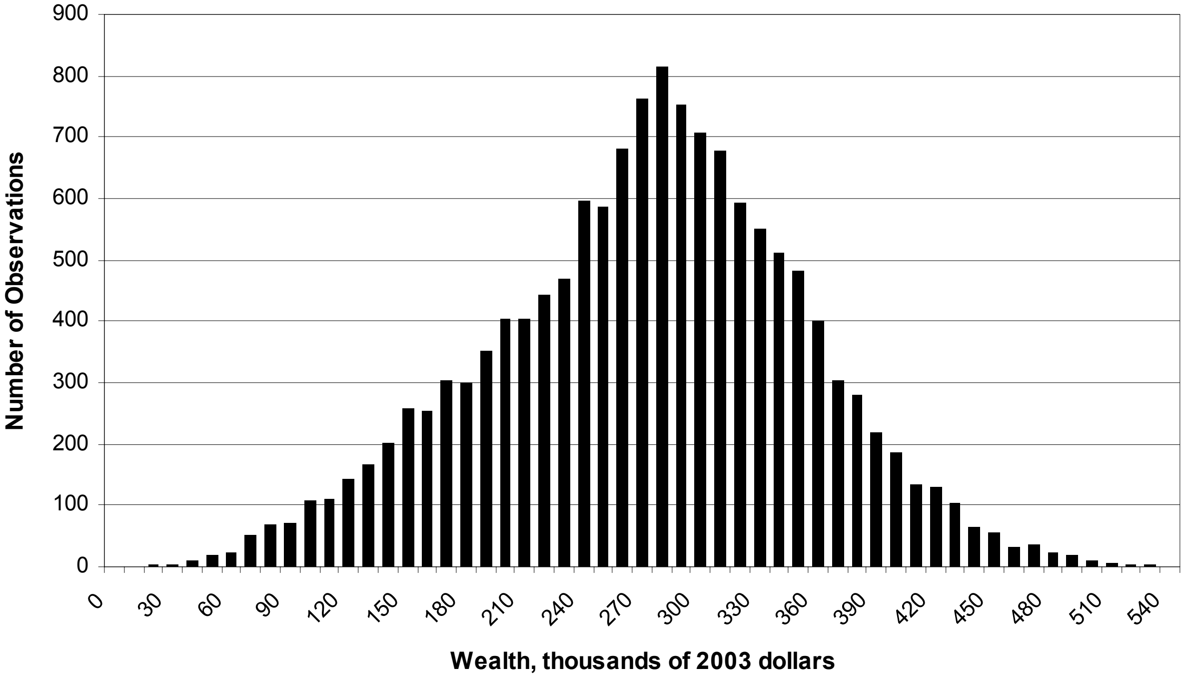
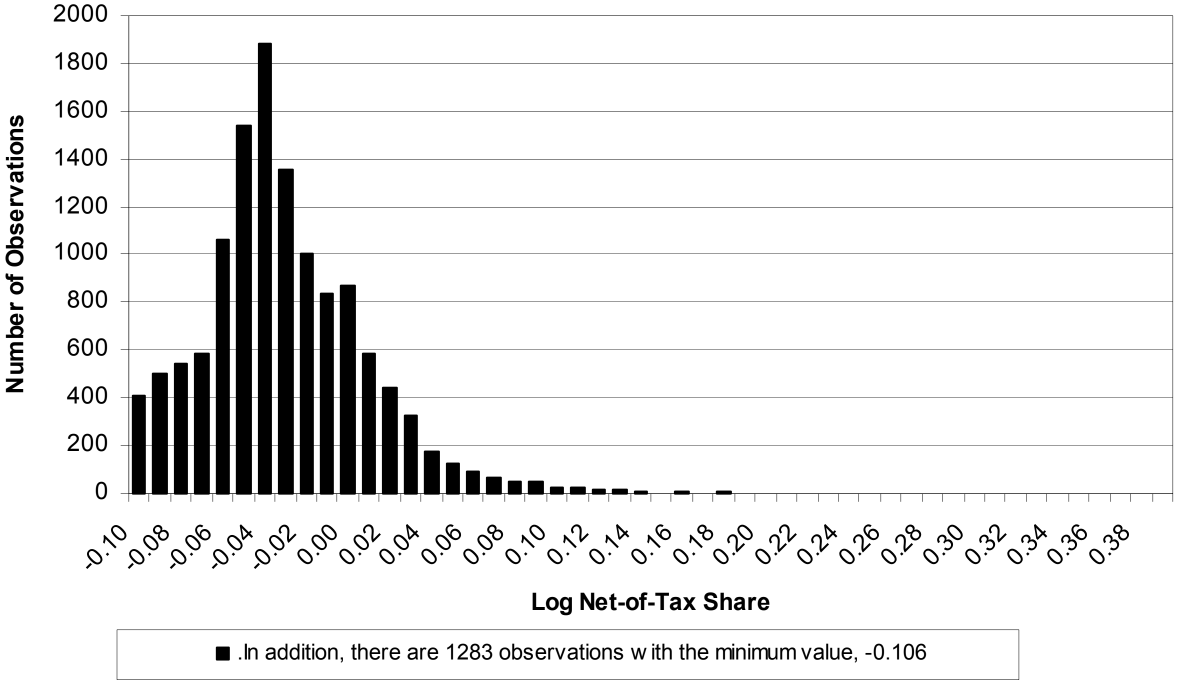
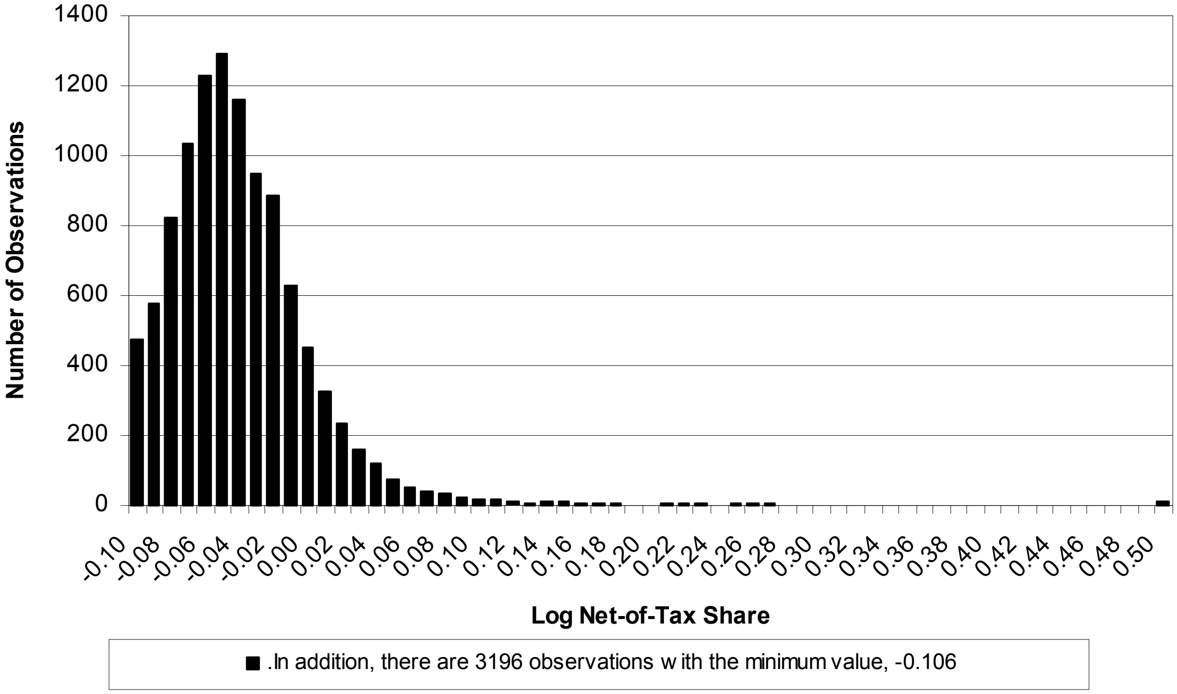


Figure 3
Social Security Net-of-Tax Share, Intensive Margin



■ .In addition, there are 1283 observations w ith the minimum value, -0.106

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



■ .In addition, there are 3196 observations w ith the minimum value, -0.106

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

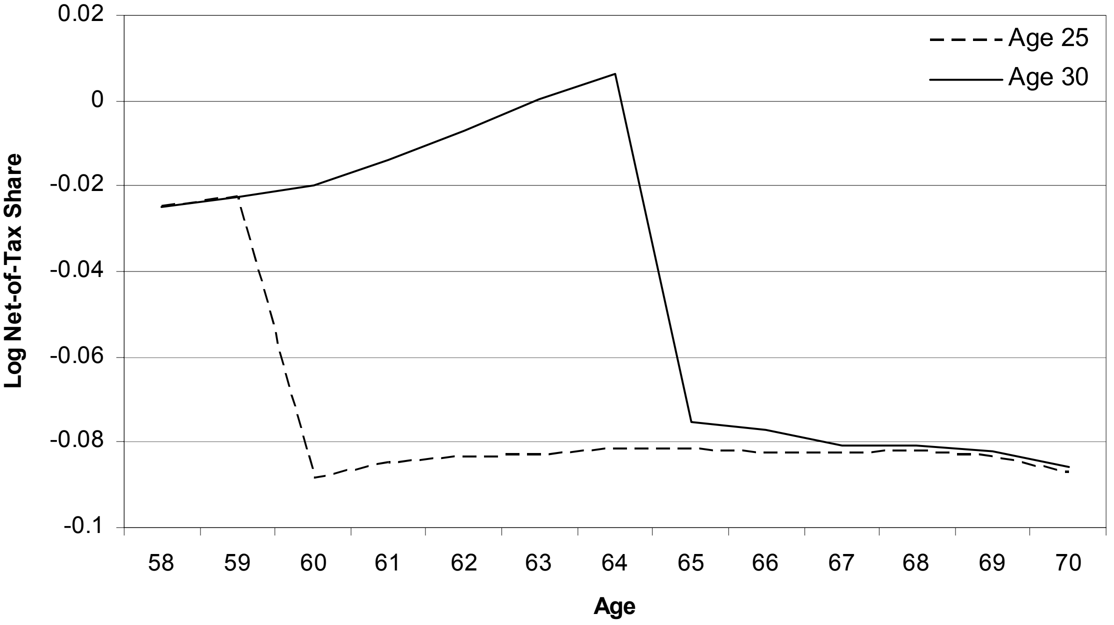


Figure 6
Social Security Net-of-Tax Share, Intensive Margin, by Age of Labor Force Entry

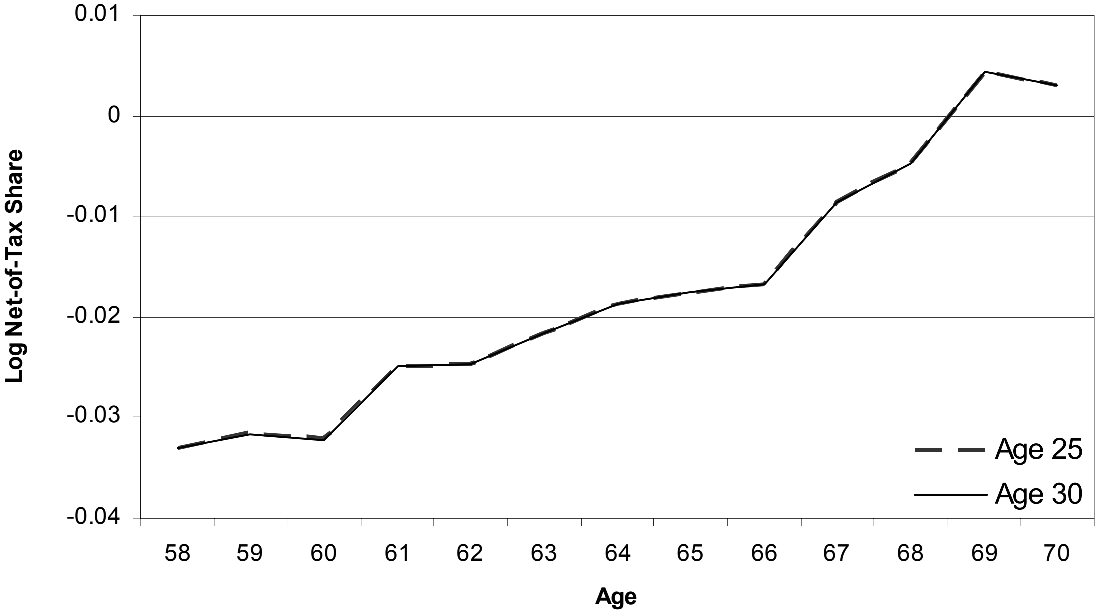


Figure 7
Social Security Net-of-Tax Share, Intensive Margin, Spouses of Retirees and Widows

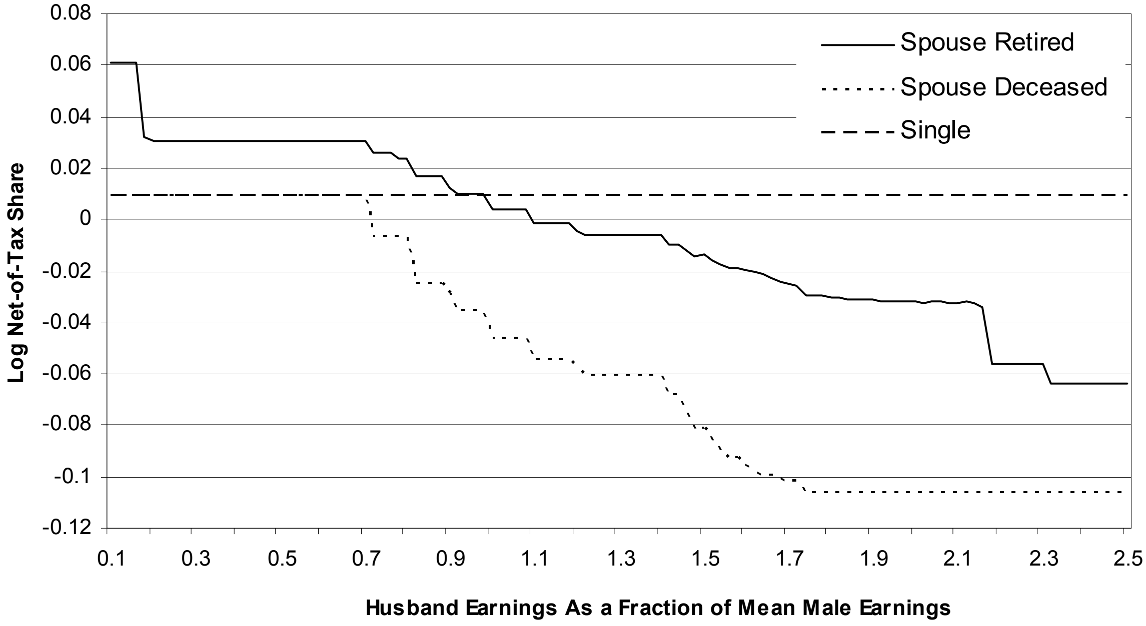


Figure 8
Social Security Net-of-Tax Share, Extensive Margin, Spouses of Retirees and Widows

Table 1
Selected Summary Statistics

	Whole Sample		Men		Women	
	Mean	(std.dev.)	Mean	(std.dev.)	Mean	(std.dev.)
Earnings	30569	(22604)	34643	(23981)	25087	(19295)
ln(Earnings)	9.398	(2.729)	9.536	(2.750)	9.212	(2.690)
Two Year Retirement Hazard	0.151	(0.358)	0.158	(0.364)	0.143	(0.350)
Hours Worked per Week	38.82	(17.58)	41.68	(18.22)	34.97	(15.89)
Dummy for Weekly Hours≥15	0.888	(0.315)	0.893	(0.309)	0.882	(0.322)
Weekly Hours if Weekly Hours≥15	43.51	(12.23)	46.58	(12.13)	39.33	(11.06)
ln(Weekly Hours) if Weekly Hours≥15	3.733	(0.288)	3.808	(0.263)	3.632	(0.291)
Fraction Retired at Age 60	0.409		0.305		0.497	
Fraction Retired at Age 65	0.759		0.686		0.821	
Fraction Retired at Age 70	0.939		0.907		0.966	
ln(SS Wealth)	12.46	(0.35)	12.48	(0.34)	12.43	(0.37)
SS Wealth	272153	(80940)	277704	(78901)	264683	(83029)
SS Wealth if Age≥62	300493	(87377)	310742	(83045)	279415	(92185)
Years of Earnings	33.10	(8.93)	37.28	(6.75)	27.49	(8.41)
Eligible for SS with Own Record	0.988	(0.110)	0.994	(0.077)	0.979	(0.142)
Age	59.40	(4.63)	60.09	(4.79)	58.47	(4.25)
Married	0.923	(0.266)	0.955	(0.207)	0.880	(0.325)
Widowed	0.035	(0.185)	0.015	(0.122)	0.063	(0.242)
Single	0.041	(0.199)	0.030	(0.169)	0.057	(0.232)
Number Person x Wave Obs.	13902		7975		5927	
Unique Persons	3971		2269		1702	

Notes: All dollars are real 2003 dollars. Fraction retired statistics are for entire HRS cohort of HRS. 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.

Table 2
Effective Social Security Net-of-Tax Share

		Social Security Wealth		Log of Intensive-margin Net-of-Tax Share (INTS)		Log of Extensive-margin Net-of-Tax Share (ENTS)	
	Number of Obs.	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 includes high school dropouts and high school graduates; high education includes 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 "Smoothened" Incentives on Control Variables

Table reports:	Retirement Regressions Dep. Variable: Smoothened ENTS			Hours Regressions Dep. Variable: Smoothened INTS			Earnings Regressions Dep. Variable: Smoothened INTS		
R^2 (top number) and Root MSE (bottom number)	Whole Sample	Men	Women	Whole Sample	Men	Women	Whole 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	0.2478	0.3922	0.6924	0.3389	0.4625	0.6145	0.3527	0.4692	0.6270
	0.0395	0.0323	0.0116	0.0305	0.0306	0.0131	0.0311	0.0314	0.0132
3. Baseline Minus Earnings History	0.7589	0.7769	0.7815	0.7135	0.8124	0.7444	0.7178	0.8179	0.7602
	0.0224	0.0197	0.0098	0.0202	0.0182	0.0107	0.0206	0.0185	0.0107
4. Baseline Minus Higher Order Terms	0.7614	0.8075	0.7994	0.7180	0.8429	0.7712	0.7218	0.8456	0.7878
	0.0223	0.0183	0.0094	0.0200	0.0166	0.0101	0.0204	0.0170	0.0100
5. Baseline Controls	0.9421	0.9368	0.8637	0.9423	0.9720	0.8446	0.9439	0.9713	0.8608
	0.0111	0.0107	0.0080	0.0092	0.0072	0.0086	0.0093	0.0075	0.0083
6. Baseline Plus Additional Interactions	0.9741	0.9635	0.9553	0.9738	0.9845	0.9323	0.9750	0.9841	0.9391
	0.0075	0.0082	0.0046	0.0062	0.0054	0.0057	0.0062	0.0056	0.0056
7. Baseline Plus Further Additional	0.9770	0.9677	0.9589	0.9772	0.9871	0.9381	0.9783	0.9868	0.9444
Interactions	0.0071	0.0077	0.0045	0.0059	0.0050	0.0056	0.0059	0.0052	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 smoothened ENTS (log of the effective Social Security Extensive-margin Net-Of-Tax Share) or smoothened 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 A1 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: R^2 (top number) and Root MSE (bottom number)	Retirement Regressions Dependent Variable: ENTS			Hours Regressions Dependent Variable:INTS			Earnings Regressions Dependent Variable: INTS		
	Whole Sample	Men	Women	Whole Sample	Men	Women	Whole 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	0.0971	0.0897	0.1609	0.1511	0.0789	0.2824	0.1484	0.0794	0.2648
	0.0471	0.0483	0.0428	0.0400	0.0441	0.0305	0.0416	0.0458	0.0323
3. Baseline Minus Earnings History	0.4090	0.4956	0.4232	0.5433	0.4745	0.6880	0.5448	0.4918	0.6717
	0.0382	0.0361	0.0357	0.0295	0.0335	0.0202	0.0305	0.0342	0.0217
4. Baseline Minus Higher Order Terms	0.4385	0.5025	0.4775	0.5535	0.4773	0.7165	0.5551	0.4946	0.6977
	0.0373	0.0359	0.0340	0.0292	0.0334	0.0193	0.0302	0.0341	0.0209
5. Baseline Controls	0.5695	0.6079	0.5651	0.6769	0.6159	0.7807	0.6736	0.6198	0.7643
	0.0330	0.0324	0.0317	0.0252	0.0293	0.0174	0.0262	0.0302	0.0189
6. Baseline Plus Additional Interactions	0.6182	0.6632	0.6037	0.7196	0.6682	0.8117	0.7154	0.6690	0.7962
	0.0313	0.0302	0.0306	0.0236	0.0275	0.0164	0.0246	0.0285	0.0178
7. Baseline Plus Further Additional Interactions	0.6381	0.6796	0.6323	0.7447	0.7001	0.8279	0.7396	0.6993	0.8111
	0.0307	0.0297	0.0298	0.0228	0.0264	0.0158	0.0238	0.0274	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 A1 contains detailed information on the exact set of control variables included in each row.

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

	Retirement Regressions Dependent Variable: Retirement			Hours Regressions Dependent Variable: ln(Hours)			Earnings Regressions Dependent Variable: ln(Earnings)		
Controls	Whole Sample	Men	Women	Whole Sample	Men	Women	Whole 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	-0.074	-0.309	-0.036	1.608	0.654	-2.125	-0.420	-1.829	-4.465
	(0.088)	(0.153)**	(0.452)	(0.161)***	(0.213)***	(0.475)***	(0.177)**	(0.275)***	(0.747)***
3. Baseline Minus Earnings History	0.058	-0.124	-0.187	1.021	0.412	-0.673	-1.045	-1.259	-3.088
	(0.150)	(0.231)	(0.487)	(0.225)***	(0.305)	(0.536)	(0.282)***	(0.431)***	(0.843)***
4. Baseline Minus Higher Order Terms	0.120	0.250	0.384	1.142	0.260	-0.391	-0.714	-1.087	-1.061
	(0.149)	(0.242)	(0.519)	(0.226)***	(0.340)	(0.557)	(0.285)**	(0.476)**	(0.871)
5. Baseline Controls	0.038	-0.249	0.669	0.448	0.843	-0.358	-0.463	0.124	-0.559
	(0.272)	(0.373)	(0.589)	(0.436)	(0.801)	(0.517)	(0.632)	(1.141)	(0.973)
6. Baseline Plus Additional Interactions	0.003	-0.238	-0.038	-0.464	-0.133	-1.000	-0.908	0.410	-1.093
	(0.393)	(0.479)	(0.998)	(0.553)	(0.910)	(0.752)	(0.866)	(1.504)	(1.255)
7. Baseline Plus Further Additional Interactions	0.116	-0.119	0.089	0.346	1.433	-0.532	-0.581	0.813	-0.550
	(0.404)	(0.493)	(0.978)	(0.566)	(0.952)	(0.733)	(0.881)	(1.581)	(1.359)

Notes: The table reports the coefficient and standard error on the "smoothened" INTS for ln(Earnings) and ln(Hours) regressions, and on "smoothened" 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 real 2003 dollars. Standard errors, clustered by individual, are in parentheses. * indicates p-value<.10; ** indicates p-value<.05; *** indicates p-value<.01. Appendix Table A1 contains detailed information on the exact set of control variables included in each row.

Table 5
Baseline Specification

Dependent Variable	Whole Sample	Men	Women	p-value Men=Women
1. Retirement				
Coefficient on ENTS	-0.174	-0.036	-0.324	0.071*
(Standard Error)	(0.081)**	(0.114)	(0.120)***	
R^2	0.224	0.246	0.242	
N	13,092	7,975	5,927	
2. ln(Hours)				
Coefficient on INTS	0.363	0.671	-0.431	0.002***
(Standard Error)	(0.158)**	(0.175)***	(0.316)	
R^2	0.271	0.241	0.293	
N	10,840	5,891	4,949	
3. ln(Earnings)				
Coefficient on INTS	0.104	-0.021	0.409	0.264
(Standard Error)	(0.219)	(0.263)	(0.430)	
R^2	0.512	0.515	0.523	
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 real 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 A1 contains detailed information on the exact set of control variables included.

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

	Retirement Regressions Dependent Variable: Retirement			Hours Regressions Dependent Variable: ln(Hours)			Earnings Regressions Dependent Variable: ln(Earnings)		
Controls	Whole Sample	Men	Women	Whole Sample	Men	Women	Whole 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)***	(0.100)	(0.125)***	(0.144)***	(0.256)***	(0.302)***	(0.367)*	(0.534)***
2. Only Basic Demographics	-0.143	-0.093	-0.235	0.944	0.684	0.345	-0.354	-1.022	0.055
	(0.069)**	(0.091)	(0.112)**	(0.133)***	(0.152)***	(0.263)	(0.134)***	(0.174)***	(0.254)
3. Baseline Minus Earnings History	-0.024	-0.065	-0.098	0.629	0.565	-0.068	0.314	-0.298	1.838
	(0.077)	(0.104)	(0.128)	(0.140)***	(0.157)***	(0.328)	(0.191)	(0.232)	(0.387)***
4. Baseline Minus Higher Order Terms	-0.121	-0.034	-0.280	0.568	0.556	-0.307	0.075	-0.306	0.930
	(0.078)	(0.101)	(0.129)**	(0.142)***	(0.159)***	(0.307)	(0.193)	(0.232)	(0.397)**
5. Baseline Controls	-0.174	-0.036	-0.324	0.363	0.671	-0.431	0.104	-0.021	0.409
	(0.081)**	(0.114)	(0.120)***	(0.158)**	(0.175)***	(0.316)	(0.219)	(0.263)	(0.430)
6. Baseline Plus Additional Interactions	-0.196	0.012	-0.357	0.259	0.579	-0.534	0.051	-0.016	0.282
	(0.085)**	(0.123)	(0.127)***	(0.162)	(0.178)***	(0.337)	(0.240)	(0.289)	(0.464)
7. Baseline Plus Further Additional Interactions	-0.173	0.009	-0.327	0.162	0.467	-0.433	0.019	-0.016	0.186
	(0.085)**	(0.123)	(0.131)**	(0.166)	(0.184)**	(0.351)	(0.246)	(0.296)	(0.482)

Notes: The table reports the coefficient and standard error on the INTS for ln(Earnings) and ln(Hours) regressions, and on the 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 real 2003 dollars. Standard errors, clustered by individual, are in parentheses. * indicates p-value<.10; ** indicates p-value<.05; *** indicates p-value<.01. Appendix Table A1 contains detailed information on the exact set of control variables included in each row.

Table 7

Robustness Checks

	Whole Sample		Men		Women	
Panel A: Retirement Regressions (ENTS)						
1. Baseline	-0.174	(0.081)**	-0.036	(0.114)	-0.324	(0.120)***
2. Smoothened Retirement Incentive Controls	-0.185	(0.083)**	-0.017	(0.118)	-0.365	(0.122)***
3. Earnings-Based Retirement Definition	-0.116	(0.081)	-0.028	(0.112)	-0.173	(0.122)
4. Retirement Definition Based on Self-reports	-0.161	(0.097)*	-0.225	(0.134)*	-0.178	(0.148)
5. Probit of Baseline Regression	-0.152	(0.075)**	-0.066	(0.104)	-0.271	(0.086)***
6. No Windsorization of ENTS	-0.124	(0.038)***	-0.061	(0.058)	-0.167	(0.066)**
7. Windsorize ENTS at 0.10 (top 1.7%)	-0.218	(0.104)**	-0.072	(0.152)	-0.366	(0.153)**
Panel B: Hours Regressions (INTS)						
1. Baseline	0.363	(0.158)**	0.671	(0.175)***	-0.431	(0.316)
2. Smoothened Retirement Incentive Controls	0.344	(0.162)**	0.657	(0.178)***	-0.412	(0.329)
3. Earnings-Based Retirement Definition	0.340	(0.167)**	0.698	(0.184)***	-0.473	(0.314)
4. Retirement Definition Based on Self-reports	0.117	(0.152)	0.397	(0.165)**	-0.541	(0.321)*
5. Hours≥10	0.298	(0.168)*	0.664	(0.183)***	-0.690	(0.354)*
6. Hours≥20	0.287	(0.150)*	0.507	(0.170)***	-0.291	(0.299)
7. Windsorize INTS at 0.10 (top 1.2%)	0.219	(0.149)	0.410	(0.164)**	-0.291	(0.310)
Panel C: Earnings Regressions (INTS)						
1. Baseline	0.104	(0.219)	-0.021	(0.263)	0.409	(0.430)
2. Smoothened Retirement Incentive Controls	0.141	(0.226)	-0.028	(0.276)	0.481	(0.432)
3. Earnings-Based Retirement Definition	0.212	(0.234)	0.093	(0.278)	0.501	(0.463)
4. Retirement Definition Based on Self-reports	-0.105	(0.227)	0.101	(0.278)	-0.286	(0.474)
5. Cut-off $5000	0.037	(0.202)	-0.055	(0.242)	0.288	(0.388)
6. Cut-off $1000	0.139	(0.237)	-0.024	(0.288)	0.632	(0.471)
7. Windsorize INTS at 0.10 (top 1.2%)	0.121	(0.256)	-0.115	(0.309)	0.697	(0.498)

Notes: Independent Variable is INTS. All dollars are real 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 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.

Appendix Table 1
Sets of Control Variables

	Interactions and higher-order terms of X_{t-1}					Components of X_{t-1}				Z_t	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
	Interaction of Previous 5 Years' Earnings	Dummies for Last Year's Earnings Percentile	Interactions of Education and Lifetime Earnings	Higher Order Terms in Earnings	Aggregates of Yearly Earnings	Earnings 2-52 Years Ago	Last Year's Earnings	Demo-graphics Relevant to SS Benefit Calculator	Spousal Variables Relevant to SS Benefit Calculator	Other Possible Determinants Of Labor Supply	Total Variables
Specification	(366 vars)	(99 vars)	(64 vars)	(8 vars)	(12 vars)	(104 vars)	(2 vars)	(102 vars)	(157 vars)	(48 vars)	
No Controls											0
Only Basic Demographics							X	Partial	Partial	Partial	20
Baseline Minus Earnings History							X	X	X	X	318
Baseline Minus Higher Order Terms					X	X	X	X	X	X	422
Baseline Controls				X	X	X	X	X	X	X	430
Baseline Plus Additional Interactions		X	X	X	X	X	X	X	X	X	571
Baseline Plus Further Additional Interactions	X	X	X	X	X	X	X	X	X	X	826

Numbers of variables in parentheses in each column include those dropped due to colinearity. Earnings are real personal SS-eligible earnings in 2003 dollars. Each year's earnings controls consist of two variables: log earnings if earnings are at least $1000 (and zero otherwise) and an indicator variable for earnings of at least $1000. Similarly, controls for other dollar-denominated variabels also consists of the log of that variable for amounts (in 2003 dollars) of at least $1000 and an indicator for the amount being at least $1000. Details on the exact control variables in each column follow.

Column 1: Interaction of previous 5 years of income denotes dummies for each separate marital status 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.

Column 2: Last year's earnings percentile are based on personal earned (SS-eligible) earnings.

Column 3: Log of lifetime earnings and education (no HS degree, HS degree, some college, college graduate) are interacted with dummies for each separate marital status by sex and by and age category cell. Lifetime earnings consist of the present discounted value of all personal SS-eligible earnings (using a 3% real discount rate).

Column 4: Higher order terms are second and third order terms of Social Security wealth, second and third order terms of lifetime income interacted with sex, and second order terms of current job tenure and total years of labor force experience.

Column 5: Aggregates of yearly earnings consist of 2 variables for Expected SS Wealth, 4 variables for lifetime earnings interacted with sex, 2 variables for spousal lifetime earnings, 2 variables for a quadratic in number of years worked in total.

Column 8: Demographics Relevant to the SS Benefit Calculator include 3 dummies for race, 4 dummies for education, 4 dummies for marital status, 7 dummies for year of observation, and 84 age*sex dummies. Race, education, and sex enter the calculator because life expectancy depends on these variables. In the "Only Basic Demographics" specification, only 13 variables are included: 4 dummies for education, 4 dummies for marital status, 3 dummies for race, and a quadratic in age.

Column 9: Spousal Variables Relevant to the SS Benefit Calculator include 1 variable for the number of years the spouse has been retired, 106 variables for earnings for each of the past 52 years, 42 age dummies, 1 variable for the age difference with the respondent, 4 education dummies, and 3 race dummies. In the "Only Basic Demographics" specification, only one spousal variable is included: spousal retirement status.

Column 10: Other Work and Earnings History Variables include 2 variabels 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 tenure at the longest job in one's career and it's square. In the "Only Basic Demographics" specification, only the 10 Census region dummies are included.