Journal of Public Economics 93 (2009) 1208–1223

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Labor supply responses to marginal Social Security benefits: Evidence from discontinuities ☆

Jeffrey B. Liebman [a], Erzo F.P. Luttmer [a,*], David G. Seif [b]

[a] Harvard Kennedy School and NBER, United States
[b] Department of Economics, Harvard University, United States

ARTICLE INFO

Article history:
Received 2 December 2008
Received in revised form 9 July 2009
Accepted 9 July 2009
Available online 22 July 2009

Keywords:
Social security
Discontinuity
Labor supply
Retirement
Tax-benefit link

ABSTRACT

A key question for Social Security reform is whether workers respond to the link on the margin between the Social Security taxes they pay and the Social Security benefits they will receive. We estimate the effects of the marginal Social Security benefits that accrue with additional earnings on three measures of labor supply: retirement age, 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 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% increase in the net-of-tax share reduces the two-year retirement hazard by a statistically significant 2.0 percentage points from a base rate of 15%. 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 0.42 and statistically significant, but we do not find a statistically significant earnings elasticity. Though we lack statistical power to estimate results within subsamples precisely, the retirement response is driven mostly by the female subsample, while the hours response comes from the male subsample.

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 respond to the Old-Age, Survivors, and Disability Insurance (OASDI) payroll tax as a pure tax, failing to recognize that the payment of Social Security taxes will increase their future Social Security benefits. With personal retirement accounts, by 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 respond to the Social Security tax as a pure tax or whether they instead realize that the *effective* marginal Social Security tax rate (the nominal tax rate minus the marginal Social Security benefit rate) is generally lower than the nominal Social Security tax rate. To our knowledge, no papers have examined whether the effective Social Security tax rate affects labor supply as measured by hours or earnings. While there is an extensive literature analyzing the effect of Social Security on retirement, Diamond and Gruber (1999) note that most of this literature ignores the effect of the marginal Social Security benefit rate (focusing instead on the effects of the level of Social Security Wealth). Moreover, as we explain later, nearly all of the papers that do account for accrual confound the retirement incentives with the benefit claiming date incentives. We instead isolate the retirement labor supply incentives. We see this, together with our examination of labor supply responses on the intensive margin (hours and earnings), as the first major contribution of this paper.

A challenge that faces all research on the incentive effects of Social Security is the concern that variation in these incentives may be correlated with unobserved determinants of labor supply. Structural models explicitly exclude such unobserved determinants from the utility function and instead focus on the question of whether the resulting preferences in combination with the Social Security rules can explain observed retirement patterns (Gustman and Steinmeier, 1986, 2005a; Rust and Phelan, 1997; Laitner and Silverman, 2008). Research that exploits variation over time in the Social Security rules can deal

☆ We thank seminar participants at the 2009 AEA meetings, the Boston Federal Reserve Bank, Harvard University, the NBER, and Wharton 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.
* Corresponding author.
E-mail address: erzo_luttmer@harvard.edu (E.F.P. Luttmer).

0047-2727/$ – see front matter
doi:10.1016/j.jpubeco.2009.07.006

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

The Social Security benefit formula contains a number of provisions that can create large variations in the effective marginal tax rate for otherwise very similar individuals (Boskin et al., 1987; Feldstein and Samwick, 1992). In particular, we exploit discontinuities generated by five provisions of the Social Security benefit formula. First, Social Security benefits depend on only the 35 highest years of indexed earnings, thus creating jumps in effective Social Security tax rates that depend on which years are included among the 35 highest years. Second, an individual receives total benefits that are the greater of either 100% of the person's own retired worker benefits or 50% of the benefit of the individual's spouse, thus creating a discontinuity in marginal benefits around the point where the Social Security benefit of one spouse is double that of the other spouse. Third, the provisions governing Social Security benefits for widows and widowers create discontinuities in marginal benefits. Fourth, kink points in the Social Security benefit schedule create discontinuities in marginal benefits, and fifth, there is a discontinuity at the point where the individual reaches sufficient quarters of earnings (generally 40, but lower for earlier cohorts) to become vested.

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

We perform our estimation using observations from the original cohort of the Health and Retirement Study (HRS)[2] after obtaining permission to link HRS observations to their administrative Social Security earnings records. 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% increase in the net-of-tax share reduces the two-year retirement hazard by a statistically significant 2.0 percentage points from a base rate of 15%. 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 0.42 and statistically significant. Though the point estimates are also positive, 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 more sensitive to changes in specification. Though we lack statistical power to estimate results within subsamples precisely, the retirement response is statistically significant in the female subsample but not in the male subsample, while the hours response is statistically significant in the male subsample but not in the female subsample. Overall, our results clearly allow us to reject the notion that labor supply is completely unaffected by the tax-benefit link in Social Security.

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 is indexed to the wage level of the year the worker turns 60 by multiplying the earnings by the ratio of average earnings in the year the worker turns 60 to the average earnings in the year in which the earnings were earned. Earnings after age 60 are not indexed. A worker's average indexed monthly earnings (AIME) are calculated by summing the indexed earnings from the worker's highest 35 years of indexed earnings (including zeros if the worker

[1] While there has been little sharp variation over time in Social Security benefit rules in the U.S., other countries have made changes in the public pension system that creates effective variation in incentive and income effects across cohorts and years. Manoli et al. (2009) use such variation in the case of Austria to identify the incentive and income effects of the public pension system on retirement decisions.

[2] The HRS is sponsored by the National Institute of Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. We use the RAND HRS Version F Data file (2006).

worked for fewer than 35 years) and then dividing by 420 (35×12). Only earnings up to the maximum taxable earnings level ($106,800 in 2009) 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% of average earnings over an initial segment, 32% over a second segment, and 15% over a final segment.

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

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

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

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

Second, the rules on spousal benefits create variation in the effective Social Security tax rate. This variation consists of non-linearities that converge to discontinuities as uncertainty about future own and spousal labor supply gets resolved. An individual receives total benefits that are the greater of 100% of the person's own retired worker benefit or 50% of the benefit of the person's spouse. When benefits are calculated, this creates a discontinuity at the point where the ratio of own to spousal PIA equals 0.5 because individuals will claim benefits on the spousal record when the ratio falls below 0.5. In this case, there is no link on the margin between own labor earnings and Social Security benefits. A similar discontinuity occurs when the PIA ratio reaches 2.0 because, at this point, the individual's spouse will also claim benefits on the individual's earnings record. When this occurs, it will discontinuously increase the tax-benefit linkage on the margin by about 50%.[4]

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

Fourth, the AIME-PIA conversion schedule contains three segments. In the first segment the PIA increases by $0.90 for every dollar increase in the AIME, in the second segment this figure is $0.32, and in the third segment it is $0.15. The kinks in the AIME-PIA conversion schedule create two discontinuities in the returns to work. First, the marginal returns to work fall by $(90-32)/90=64\%$ at the first kink point and by $(32-15)/32=53\%$ at the second kink point. For those who still face uncertainty about which segment they will be on, the returns to work are a weighted average of the returns to work at each of the segments, weighted by the probabilities of ending up on each. This uncertainty about future earnings turns the discontinuities into non-linearities in the returns to work around the earnings levels corresponding to the kink points.

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

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

[3] Social Security discontinuities not studied in this paper include: (1) Income taxation of benefits — The 1993 Omnibus Budget Reconciliation Act increased the fraction of Social Security benefits subject to income taxation for higher-income individuals, thus increasing effective Social Security tax rates for those individuals. (2) Divorce — Eligibility for spousal benefits upon divorce is limited to individuals who were married for at least 10 years, thus creating a discontinuity in marginal Social Security benefits at 10 years of marriage for individuals who might claim spousal or widow benefits. (3) Remarriage — Individuals lose eligibility for spousal benefits based on an ex-spouse upon remarriage prior to age 60, thus creating jumps in marginal Social Security benefits upon remarriage for the subgroup of individuals who would have claimed benefits based on an ex-spouse's earnings. (4) The Windfall Elimination Provision — This provision places workers who receive a government pension from a job in a sector not covered by Social Security on a different benefit schedule. (5) Changes in state “double-dipping” laws — These laws prevent workers from receiving state pensions from SS-ineligible government work if they are claiming any Social Security, thus effectively forcing many workers not to take Social Security benefits. (6) The “Special Minimum PIA” — This creates variation in effective marginal Social Security benefit rates for workers with similar lifetime earnings but with different year-by-year earnings histories. (7) Children's benefits — Minor children of retirees are eligible to receive 50% of the retiree's benefits, which creates variation in effective marginal Social Security benefits based on the age difference between the parent and child.

[4] The increase is exactly 50% if the individual and the spouse are the same age, have the same life expectancy, and retire at the FRA. In other cases, differences in life expectancy and adjustments for age of benefit take-up can cause this increase to be somewhat larger or smaller than 50%.

[5] As with spousal benefits, adjustments for age of benefit take-up may result in slightly different values.

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

If individuals had perfect foresight, we could use a standard regression discontinuity design to exploit the discontinuities generated by the five provisions in the Social Security benefit rules we discussed above (e.g., see Hahn et al., 2001 for the standard regression discontinuity design). In particular, we could calculate the present discounted value of all future Social Security benefit payments for person i and his or her spouse: SSW_{it} $(\boldsymbol{X}_{i,t-1}, \boldsymbol{X}^{+}_{i,t-1})$, where $(\boldsymbol{X}_{i,t-1}, \boldsymbol{X}^{+}_{i,t-1})$ is the vector of individual characteristics (including own and spousal earnings) that determine Social Security benefit payments. This vector consists of a component, $\boldsymbol{X}_{i,t-1}$, that is known at time $t-1$, and a component, $\boldsymbol{X}^{+}_{i,t-1}$, that is not yet known at that time (except under perfect foresight). The person would face an effective Social Security tax of:

$$\tau_{it}^{\text{effective}}\left(\boldsymbol{X}_{i,t-1},\boldsymbol{X}^{+}_{i,t-1}\right) = \tau_t^{\text{nominal}} - \partial SSW_{it}\left(\boldsymbol{X}_{i,t-1},\boldsymbol{X}^{+}_{i,t-1}\right) / \partial y_{it}, \quad (1)$$

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}:

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

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_{it} is an error term. The functional form of the net-of-tax share, $1-\tau_{it}^{\text{effective}}$, 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:

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

Due to the expectation operator, $E[.]$, many discontinuities in the effective marginal tax rate turn into non-linearities. These non-linearities would be fully absorbed by $f(.)$ if we were to allow $f(.)$ to be an arbitrarily flexible function of $\boldsymbol{X}_{i,t-1}$, and, as a result, the labor supply response to the net-of-tax share would no longer be identified. This creates a dilemma. On the one hand, we want $f(.)$ to be sufficiently flexible to capture any relation between past determinants of the expected effective Social Security tax rate ($\boldsymbol{X}_{i,t-1}$) and unobserved determinants of labor supply (ε_{it}). On the other hand, we require sufficient remaining variation in the effective marginal tax rate to identify the labor supply effects. The key to our methodology is the creation of a criterion that allows us to determine whether the control function $f(.)$ is sufficiently flexible.

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

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

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

In these regressions, the effect of the smoothed effective tax rate on labor supply is purely identified off of non-linearities in the Social Security benefit schedule such as the progressive nature of the AIME-PIA schedule (now modeled as a quadratic relationship) or the fact that the present discounted value of benefits increases as individuals age (since older individuals are closer to receiving benefit payments than younger people are). Even though some of this variation may be valid, we are not comfortable using it because many of these non-linearities may be gradual and could plausibly be correlated with unobserved determinants of labor supply. To ensure that none of this variation drives our main estimates (from Eq. (3)), we increase the flexibility of the functional form of the control function $f(.)$ until the estimate of α in the auxiliary regressions (Eq. (4)) becomes completely insignificant. We then use that functional form for the control function in the main regression.

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

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

3. Data and empirical implementation

3.1. Data

We perform our estimation using data from the original cohort of the Health and Retirement Study (HRS), a longitudinal survey that can be linked to Social Security earnings records. This cohort consists of individuals born between 1931 and 1941 as well as their spouses. Individuals were first interviewed in 1992 and have been re-interviewed every two years. Our data extend through the seventh wave of the HRS, which was conducted in 2004. In total, the original cohort of the HRS includes 12,582 individuals who were interviewed at least once.

Table 1
Selected summary statistics.

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

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

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

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

The HRS survey data also contain the two other dependent variables for our regressions: earnings and hours worked per week. The first of these is self-reported, with answers corresponding to the previous year. Our hours worked variable is the sum of the usual hours per week individuals report working at their primary and secondary jobs measured at the time of the survey. In addition, the HRS data contain information that allows us to create an extensive set of control variables. Data are collected semi-annually in even years, but financial variables other than household wealth 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 (about one-third of potential observations). We also exclude individuals who were already retired before the initial wave of the HRS or who had very weak past labor force attachment (about 17% of potential observations). In addition, we exclude widowed, separated, and divorced individuals in cases for which we have insufficient information about their former spouses to calculate benefits (about 11% of potential observations). Furthermore, we exclude anyone who reports ever having been disabled in the HRS (about 6% of potential observations), as disability changes Social Security incentives in ways that we do not have sufficient information to model correctly. Other sample restrictions result in much smaller numbers of dropped observations, leaving us with a sample of 3971 individuals (2269 men and 1702 women) out of the 12,582 individuals in the original HRS cohort. See Online Appendix Table 1 for a full list of sample selection criteria. For our analysis of hours and earnings, slightly fewer observations are included, as described in the Appendix Table 1.

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

Table 1 shows summary statistics for the key variables in our data. In each two-year wave, an individual has approximately a 15.1% chance of retiring, and this hazard rate does not vary significantly by sex. Conditional on working, the average male respondent works almost 42 h per week while the average female respondent works 35 h per week. Mean Social Security Wealth discounted at a 3% real rate is $272,153. Nearly all sample members, male and female, have had sufficient earnings histories to be eligible for Social Security benefits as retired workers. In constructing our sample, we dropped most of the individuals who were unmarried at the time of the first wave of the HRS, so 92% of the person-year observations in our sample came from married individuals. The average age is 60 for men and 58 for women. On average, men have had earnings in 37 prior years and women in 27 prior years.

Fig. 1 shows annualized two-year retirement hazard rates by gender. The figure shows that there is a considerable age range within which retirement hazard rates are substantial. We find that for both

[6] Social Security benefits for individuals in our sample do not depend on earnings from years prior to 1951.

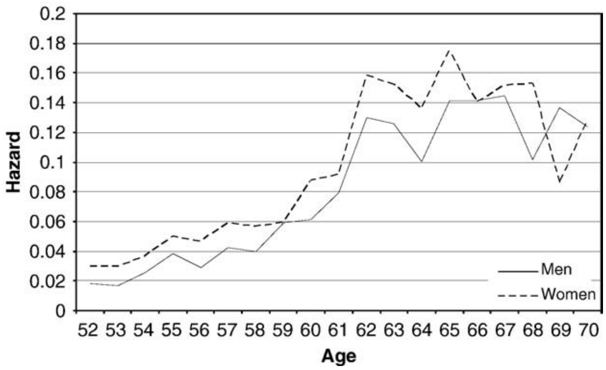

Fig. 1. Annualized two-year retirement hazard.

men and women the retirement hazard rate more than doubles from 6% to above 12% between ages 60 and 62 and then remains relatively constant thereafter.

3.2. Calculating expected Social Security Wealth

We define the effective Social Security tax rate as the nominal Social Security tax rate (10.6%)[7] minus the expected Social Security marginal benefit rate, where this benefit rate is defined as the marginal effect of current labor supply on expected Social Security Wealth. Thus, the calculation of Social Security Wealth is a key element of our analysis. We define Social Security Wealth as the expected present discounted value of all payments from the Social Security Administration to the individual and, if the individual is married, to his or her spouse. Future Social Security benefits are calculated using the current Social Security benefit 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, early retirement benefit reductions, delayed retirement credits, and the vesting rule based on quarters of earnings.[8] We model the benefits workers can claim on their own earnings record and any additional benefits they are entitled to based upon the record of their living or deceased spouse. We update the benefit calculation (i) when the individual first claims benefits, (ii) when the individual first becomes eligible to claim benefits on the spousal record, (iii) when the spouse dies, or (iv) if claiming widow benefits, when the individual first becomes eligible to claim benefits on his or her own record.[9] Further details of the benefit calculation are spelled out in Online Appendix B.

Future Social Security benefits are a non-linear function of (i) own year of birth, (ii) spousal year of birth, (iii) own earnings history, (iv) spousal earnings history, (v) future own earnings, (vi) future spousal earnings, (vii) year of own death, (viii) year of spousal death, (ix) year in which the individual starts claiming benefits, and (x) year in which the spouse starts claiming benefits. Year of birth and earnings history are known, but the remaining eight variables are generally stochastic. Thus, future Social Security benefits are an expectation with respect to eight variables. We reduce the dimensionality of this expectation by specifying the year of benefit take-up as a function of age and year of retirement (so, conditional on age and year of retirement, year of benefit take-up is not stochastic and we do not need to take an expectation over it).[10] However, as explained in more detail in Section 3.3, when computing Social Security Wealth to calculate retirement incentives, we keep the year of own benefit take-up constant as we vary the year of own retirement. To reduce the computational burden, we further assume that retirement occurs no later than at age 80 and that death occurs no later than at age 100.

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

3.3. The expected effective Social Security tax rate

The Social Security benefit schedule generally has different incentive effects on the extensive and intensive margins of labor supply. Following the convention in public economics, we measure the incentive effect by the log of the net-of-tax share, $\ln(1-\tau)$, where τ is the effective marginal Social Security tax.

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:

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

where SSW_{it} denotes the individual's expected Social Security Wealth at time t, and $\hat{y}_{it}$ denotes the person's predicted pre-Social Security tax earnings for year t.[12] Because *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.

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

[8] We do not model the Special Minimum PIA because, by our calculation, it would apply to less that 0.1% 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 so. Since we exclude individuals with more than 10 years of non-FICA-covered work, these provisions would apply to very few of the observations that remain. In order to model them, we would need more information than is available in the HRS. 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 apply to very few individuals in our sample.

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

[10] In particular, we assume the individual starts claiming benefits in the year of retirement with two exceptions: (i) if the individual retires before the early retirement age, we assume that the individual starts claiming benefits at age 62 (even if widowed and eligible at age 60), and (ii) we assume those who are not retired at age 70 will nevertheless start claiming benefits then (there is never any benefit to delaying claiming benefits beyond age 70 because the delayed retirement credit does not increase after age 70).

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

[12] The 10.6% 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%.

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. We define the expected effective Social Security extensive-margin net-of-tax share (ENTS) for individual i in year t as:

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

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).[13] This separation of the retirement incentives from the benefit claiming incentives contrasts 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 people who continue working for one more year also delay claiming for one more year.[14] 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 the response to the link between the work decision (rather than the claiming decision) and the level of future benefits.

4. Results

4.1. Effective Social Security net-of-tax shares

Before estimating the labor supply response to incremental Social Security benefits, we first present our estimates of Social Security Wealth and the corresponding intensive-margin and extensive-margin net-of-tax shares. We do this for two reasons. First, the size and variation in the incentives implicit in the Social Security rules are of interest in and of themselves as they inform how benefit rules could be restructured to reduce the size and variation of distortions. Indeed, these incentives are the focus of a number of papers in the literature. See, for example, Feldstein and Samwick (1992), Butrica et al. (2006), Goda (2007), Sabelhaus (2007), and Goda et al. (2009). Second, we want to document the variation in the incentives. If the variation in the estimated incentives corresponds to what we would expect given the Social Security rules, we can be more confident that our calculated incentives are correct.[15]

Fig. 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.[16] Future benefits are discounted to the present using a 3% real discount rate. Median Social Security Wealth is \$269,000 while the Social Security Wealth of 90% of our sample ranges between \$0 and \$360,578. These values are in line with those found in the literature.[17] The second and third columns of Table 2 show the mean and standard deviation of Social Security Wealth by demographic subgroup. As expected, Social Security Wealth is higher for married individuals than for widowed or single individuals and increases with work history, lifetime earnings, and education.

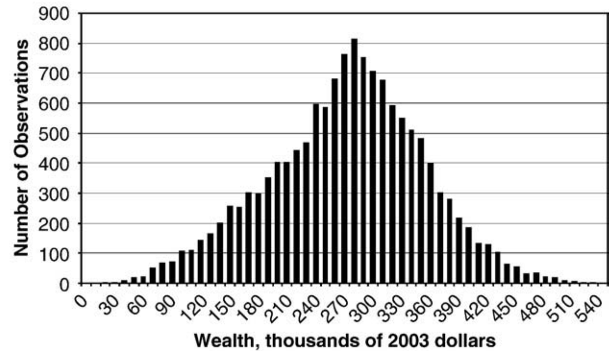

Fig. 2. Expected family social security wealth.

Fig. 3 shows the distribution of the log of the effective Social Security intensive-margin net-of-tax share (INTS), as defined by Eq. (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%, and the log of their net-of-tax share is $\ln(1-0.106/1.053) = -0.106$. Because additional earnings can never reduce expected Social Security benefits, this is also equal to the minimum of the log of the net-of-tax share. The mean INTS is -0.037, which corresponds to a net marginal Social Security tax rate of 3.8%. Thus, on average, the effective Social Security tax is 6.8 percentage points lower than the nominal tax due to the tax-benefit linkage. However, the tax-benefit linkage varies tremendously and is highly right skewed. Whereas 20% of person-years have virtually no tax-benefit linkage (INTS < -0.10), the tax-benefit is sufficiently strong for 18% 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 show the mean and standard deviation of INTS by demographic subgroup. Work incentives are lower for women than for men because women are more likely to claim off of the records of their spouses. Among men, work incentives are stronger for those with shorter work histories and lower lifetime earnings. These effects are driven by the progressive nature of the Social Security benefit structure and the 35-year rule, giving those with lower earnings and fewer working years 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 scenario makes it much more likely that they will claim based on their own record. This also explains why, despite the progressive nature of the benefit schedule, work incentives are generally relatively weak for women with short earnings histories or low lifetime earnings—there is no tax-benefit link if they claim on their spouses' records.

---

13 We acknowledge, but do not model, the option value in the decision not to retire, as highlighted by Stock and Wise (1990). Option value is more important in models that assume that the age of benefit take-up coincides with the age of retirement because in those settings the option value not only includes the option of adding more years of earnings to the earnings history but also the option of further delaying benefit take-up. Nevertheless, an interesting extension would be to take a peak-value approach as in Coile and Gruber (2007) since this would make it possible to account in part for the fact that returns to work in later years might affect the decision whether or not to retire in the current year.

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

15 We verified that our calculator of Social Security benefits yielded the identical level of benefits as the calculator provided by the Social Security Administration (www.ssa.gov/retire2/AnypiaApplet.html). 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.

16 We have no valid observations older than age 80. All such individuals in our data were either retired in the first wave or were born prior to 1920, making them subject to different Social Security benefit rules.

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

Table 2
Effective Social Security net-of-tax shares.

	Number of obs.	Social Security wealth		Log of intensive-margin net-of-tax share (INTS)		Log of extensive-margin net-of-tax share (ENTS)	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Panel A: whole sample							
All observations	13,902	272,153	(80,940)	−0.037	(0.045)	−0.054	(0.050)
All men	7975	277,704	(78,901)	−0.029	(0.047)	−0.048	(0.051)
All women	5927	264,683	(78,901)	−0.048	(0.039)	−0.063	(0.051)
Panel B: men only							
By work history							
35 years of work or fewer	2335	224,758	(74,636)	−0.008	(0.047)	−0.007	(0.060)
More than 35 years of work	5640	299,625	(69,685)	−0.037	(0.044)	−0.064	(0.034)
By marital status							
Married	7619	284,469	(73,481)	−0.028	(0.047)	−0.047	(0.051)
Widowed	120	157,300	(40,547)	−0.058	(0.035)	−0.073	(0.031)
Single	236	120,521	(40,883)	−0.045	(0.036)	−0.049	(0.046)
By education:							
Low education	4446	253,283	(71,258)	−0.026	(0.050)	−0.052	(0.052)
High education	3529	308,471	(77,296)	−0.032	(0.043)	−0.043	(0.048)
By lifetime earnings:							
Lifetime earnings < median	3997	234,182	(70,219)	−0.008	(0.050)	−0.028	(0.059)
Lifetime earnings ≥ median	3978	321,435	(60,918)	−0.050	(0.032)	−0.067	(0.028)
By ratio of own to spousal lifetime earnings:							
Ratio of earnings 1st quartile	1905	287,731	(84,527)	−0.029	(0.043)	−0.049	(0.048)
Ratio of earnings 2nd quartile	1905	292,073	(69,358)	−0.033	(0.043)	−0.051	(0.047)
Ratio of earnings 3rd quartile	1905	284,912	(66,126)	−0.025	(0.049)	−0.045	(0.053)
Ratio of earnings 4th quartile	1904	273,157	(71,285)	−0.024	(0.052)	−0.045	(0.055)
Panel C: women only							
By work history:							
35 years of work or fewer	4869	259,852	(81,536)	−0.049	(0.039)	−0.062	(0.049)
More than 35 years of work	1058	286,917	(86,193)	−0.044	(0.038)	−0.068	(0.032)
By marital status:							
Married	5217	280,376	(73,935)	−0.048	(0.037)	−0.064	(0.044)
Widowed	371	153,622	(38,933)	−0.067	(0.048)	−0.084	(0.055)
Single	339	144,721	(55,446)	−0.026	(0.043)	−0.028	(0.058)
By education:							
Low education	3484	245,974	(76,626)	−0.052	(0.038)	−0.068	(0.044)
High education	2443	291,365	(84,515)	−0.043	(0.039)	−0.056	(0.050)
By lifetime earnings:							
Lifetime earnings < median	2971	238,529	(77,313)	−0.055	(0.044)	−0.072	(0.055)
Lifetime earnings ≥ median	2956	290,970	(80,238)	−0.041	(0.031)	−0.055	(0.034)
By ratio of own to spousal lifetime earnings:							
Ratio of earnings 1st quartile	1305	274,723	(60,057)	−0.079	(0.027)	−0.099	(0.024)
Ratio of earnings 2nd quartile	1304	286,963	(68,235)	−0.050	(0.027)	−0.071	(0.037)
Ratio of earnings 3rd quartile	1304	291,942	(75,412)	−0.038	(0.029)	−0.050	(0.032)
Ratio of earnings 4th quartile	1304	267,880	(86,973)	−0.027	(0.040)	−0.036	(0.050)

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

Fig. 4 shows the distribution of the log of the effective Social Security extensive-margin net-of-tax share (ENTS), as defined by Eq. (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. This minimum is reached for 33% of the sample, and occurs for a variety of reasons. For example, current predicted indexed earnings might not be among the 35 highest annual indexed earnings, the person might be sure to claim spousal or widow benefits, or the individual might not be vested.

At the other extreme, 0.2% of observations have tax-benefit linkages that are so strong that their ENTS exceeds 0.5 (i.e., they receive an effective subsidy of about 68%). All of these cases occur because one additional year of earnings will give the individual sufficient quarters of earnings to qualify for receiving Social Security benefits. For example, we have one individual whose predicted earnings of $3467 give this person a total of 40 quarters of earnings, thereby qualifying this person for $112,308 in Social Security Wealth. This translates into an effective subsidy of 3239% on earnings, or an ENTS of 3.51. 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 observations that reach the required number of quarters of earnings in that year to qualify for Social Security. To avoid this, we topcode ENTS at 0.50, which is slightly above the

[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.

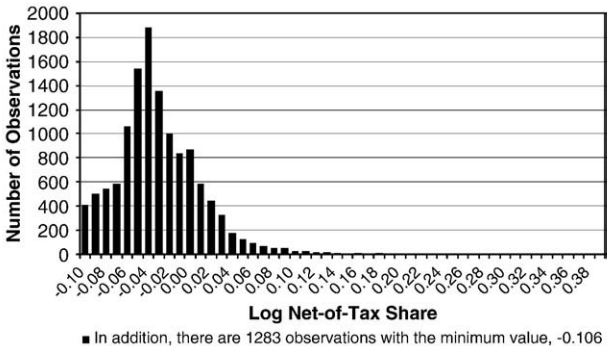

Fig. 3. Social Security net-of-tax share, intensive margin.

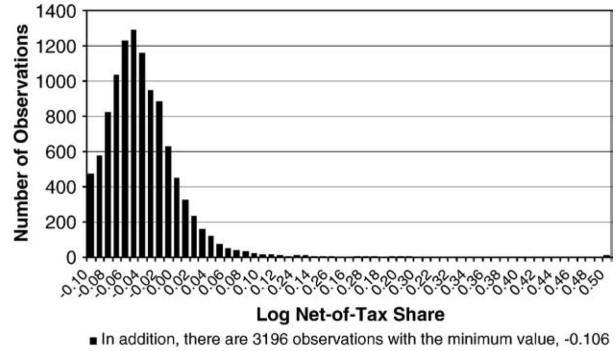

Fig. 4. Social Security net-of-tax share, extensive margin.

highest value of ENTS achieved by someone with more than 40 quarters of earnings. Even after topcoding, the distribution of ENTS remains right-skewed with a median of −0.060 and a mean of −0.054. Ninety percent of observations in our sample have an ENTS that is negative and thus face a work disincentive. Columns 6 and 7 of Table 2 show the mean and standard deviation of ENTS by demographic subgroup. Even though the INTS and ENTS can differ strongly for a given individual in a particular year, the overall variation in ENTS is similar to the variation in INTS when broken down by broad population subgroups as in Table 2.

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

Figs. 5–8 illustrate how provisions of the Social Security benefit rules 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.

Fig. 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 an earlier 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.

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

The provision that allows one to choose between claiming benefits based on 100% of own PIA or 50% of spousal PIA creates discontinuities-in-the-limit in both the intensive-margin incentives (INTS) and the extensive-margin incentives (ENTS). Consider a hypothetical 63-year-old woman who started working at age 30 and whose earnings are equal to the mean earnings profile for women. Her husband is also 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 Fig. 7 plots the INTS of this woman as a function of x, the fraction of age-specific male mean earnings that her husband earned. We see three drops in the INTS, two of which are quite sudden. The first drop occurs around x equal to 0.2 and is associated with her husband's PIA rising above 50% of her expected PIA so that he chooses to claim benefits on his own rather than her record, while she continues to claim benefits on her own record. The second drop occurs more gradually between x equal to 0.7 and 1.7 because over this range it becomes increasingly likely that it will be beneficial for her to claim benefits on his record if she outlives him, and at the same time, that he will choose to claim benefits on his own record should he outlive her. In other words, uncertainty about her future labor supply has turned the discontinuity associated with claiming widow/widower benefits into a non-linearity. Finally, the third drop occurs at x equal to 2.2%, 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 while her effective Social Security tax 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 0.7, where she switches from claiming on her own record to claiming on her spousal record. The decline is only a discontinuity-in-the-limit because there is still uncertainty about her ultimate PIA due to uncertainty about her retirement date. Finally, the long dashes show the INTS for a hypothetical woman who is the same except that she has always been single. Fig. 8 plots the extensive-margin incentives (ENTS) for the same three hypothetical women, and we observe a very similar pattern of drops in incentives, except that the non-linearities are true discontinuities. We observe true discontinuities in extensive-margin incentives because these incentives measure the benefit of retiring in the current year relative to retiring in the next year, and therefore the women's PIAs at the retirement dates considered are known. Moreover, in this example, the husband was already retired or deceased, so his PIA was also known.

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

To ensure that labor supply incentives are identified 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 $f(\boldsymbol{X}_{i,t-1}, \boldsymbol{\gamma})$ that is sufficiently flexible to absorb the remaining variation in the ENTS and INTS.

[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 zero, meaning her effective tax rate never quite reaches 10.6%. However, if she were a somewhat lower earner (e.g. if she had earned each year only 75% of the mean earnings profile for women), she would never claim on her own record for sufficiently high values of x, and her tax-benefit linkage would reach zero.

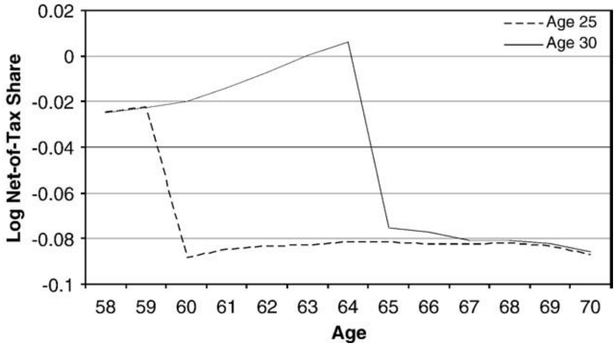

Fig. 5. Social Security net-of-tax share, extensive margin, by age of labor force entry.

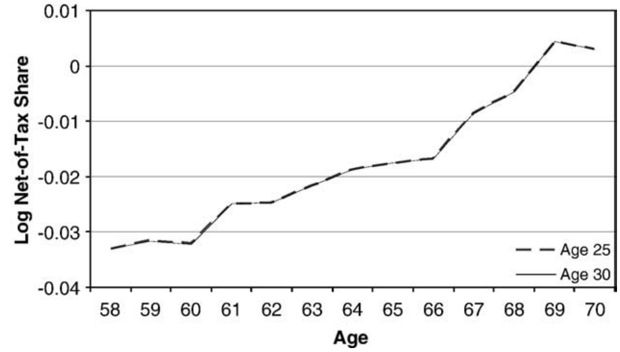

Fig. 6. Social Security net-of-tax share, intensive margin, by age of labor force entry.

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

Our baseline control function $f(\boldsymbol{X}_{i,t-1}, \boldsymbol{\beta})$ consists of a linear combination (with weights $\boldsymbol{\beta}$) 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 (married, widowed, never married), for own and spousal education (high school dropout, high school graduate, some college, bachelor's degree or more), for own and spousal race-ethnicity (non-Hispanic black, non-Hispanic white, Hispanic, other), for the retirement status of the individual's spouse, and for the calendar year in which the observation takes place. Further, the linear combination includes a cubic polynomial in the log of 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, a cubic polynomial in the number of years in which own Social Security earnings exceeded $1000, and the same cubic polynomial for the spouse. In total, $f(\boldsymbol{X}_{i,t-1}, \boldsymbol{\beta})$ consists of 386 terms.

The baseline vector of additional controls, $\boldsymbol{Z}_{it}$, consists of the log of household assets, a cubic polynomial for job tenure at current job, and dummies for veteran status, for being born in the U.S., for longest industry of employment (13 dummies), for longest occupation (17 dummies), for the 10 Census regions, for self-reported health status interacted with sex (11 dummies), and for seven chronic health conditions interacted with sex (21 dummies). In total, $\boldsymbol{Z}_{it}$ consists of 78 terms, for a total of 464 terms between $\boldsymbol{X}_{i,t-1}$ and $\boldsymbol{Z}_{it}$. We do not include the marginal labor supply incentives from private pensions as a control variable because this information is missing for a large fraction of respondents. While this would potentially be a useful control variable, its effects are unlikely to be correlated with the discontinuities-in-the-limit that we use to identify our estimates. Moreover, as Coile and Gruber (2007) show, the estimated incentive effects from Social Security are very insensitive to the inclusion of pension incentive variables.

The first row of Table 3a shows the raw variation in the smoothed ENTS and INTS in the samples that will be used to run the labor supply regressions. The standard deviation of the smoothed ENTS is 4.55 percentage points in the sample for the retirement regressions while the standard deviation for the smoothed INTS is 3.75 percentage points in the sample for the hours regressions and 3.86 percentage points in the sample for the earnings regressions. The fifth row of Table 3a shows that the baseline controls explain about 95% of the variation in the smoothed incentives and reduce the RMSE of the incentives to about 0.9 to 1.1 percentage points. Rows 2 to 4 show which components of the baseline control variables contribute the most to explaining the smoothed incentives. (Online Appendix Table 2 contains exact definitions of the control variables in each row.) Basic demographics only explain about 25–35%, and unless either income history or higher-order terms are included, the R^2 does not rise above about 80%. Row 6 adds an additional 163 interaction and higher-order terms to the baseline regression, which increases the explanatory power to about 97%. Row 7 shows that adding a further 320 interaction and higher-order terms to the specification from row 6 only produces a very slight increase in explanatory power.

Table 3b runs the same set of regressions on the true incentives, which include variation due to the various discontinuities-in-the-limit. As expected, the RMSE is higher for the true incentives than the smoothed ones because of this additional variation. Conversely, the R^2 is lower in all specifications. The baseline set of controls is only able to explain about 58% of the variation in the extensive-margin incentives and about 68% 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% of the variation in the ENTS and about 26% of the variation in the INTS cannot be explained. We conclude that this unexplained variation is primarily due to the discontinuities-in-the limit resulting from the five provisions

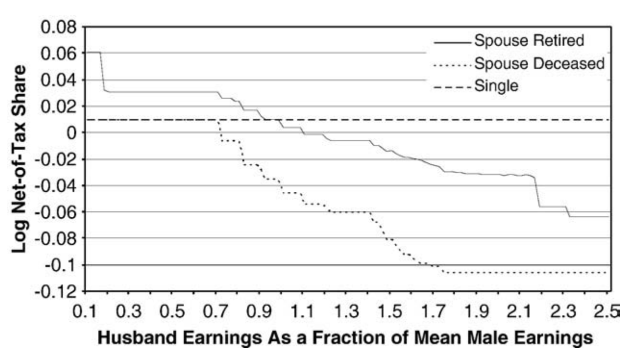

Fig. 7. Social Security net-of-tax share, intensive margin, spouses of retirees and widows.

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

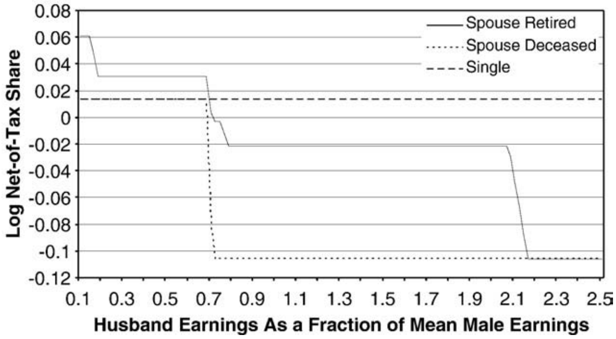

Fig. 8. Social Security net-of-tax share, extensive margin, spouses of retirees and widows.

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 is not due to these five provisions but still leaves sufficient variation in the true incentives to identify labor supply incentives. The specifications of rows 6 and 7 are more conservative because they increase the fraction of unwanted variation absorbed from about 93% to almost 99%, but they do this at the cost of possibly also absorbing some of the valid variation, namely non-linearities caused by the five provisions that have not yet converged to discontinuities.

Table 4 tests whether the control function passes our criterion of absorbing sufficient variation to render all estimates on the smoothed incentive variables insignificant. In this table, we show regressions of a measure of labor supply (retirement, hours, or earnings) on the smoothed incentive measure and on various specifications of the control function. We find that the baseline control function (in row 5), as well as the more extensive ones (rows 6 and 7), pass the criterion—none of the estimated incentive effects is significant. Less flexible control functions (rows 2–4), however, 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 five provisions in the Social Security benefit rules described in Section 2.2. Since these discontinuities-in-the-limit are driven by specific quirks in the benefit rules, we consider it highly unlikely that there could be omitted variables that are correlated with both these discontinuities-in-the-limit and with labor supply. We therefore regard our estimates as plausibly causal.

The first row of Table 5 regresses a retirement dummy on the log of the expected effective Social Security extensive-margin net-of-tax share (ENTS) as well as the 464 baseline control variables. The first column shows the estimates for the full sample while the next two columns separate the estimates by gender. For the full sample, we find that a higher net-of-tax share has a statistically significant negative effect on the retirement probability. In particular, an increase in the net-of-tax share of 0.10 (about two standard deviations) reduces the two-year retirement hazard by about 2.0 percentage points on a base of 15.1, or by about 14%. Thus, the effect is not only statistically significant but also economically meaningful. The next two columns show that this estimate appears to be driven by the subsample of women. For men, we do not find a significant effect of ENTS, but the confidence interval is sufficiently wide that we cannot rule out that the retirement response for men is as large as the point estimate for the whole sample. The hypothesis that men and women react the same way to the extensive-margin incentives cannot be rejected (p-value: 0.143).

The second and third rows of the table show the responses on the intensive margin. In the second row, we use the log of usual weekly hours as the dependent variable, while the dependent variable in the third row is the log of annual earnings. Since the expected effective Social Security intensive-margin net-of-tax share (INTS) is also measured in logs, the estimates can all be interpreted as labor supply elasticities. As we are interested in measuring labor supply responses along the intensive margin, we require that the respondents be in the labor market and subject at the margin to the OASDI tax to be included in the regression. We therefore impose a requirement of at least 15 h of labor supply per week in the hours regression or annual earnings of at least \$2500 in the earnings regression. We also require that predicted earnings are below the maximum taxable earnings for the OASDI tax. We estimate a statistically significant labor supply elasticity of 0.42 in whole sample when labor supply is measured by hours.[21] 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 0.28. We reject the hypothesis that the hours elasticity is the same for men and women. We are not sure why the estimated hours elasticity is significantly larger for men than for women. Perhaps the selected sample of women who are working around the age of 60 is truly less elastic in their hours response than men. Alternatively, different provisions in the Social Security benefit rules may be responsible for the variation in incentives for women than for men, and the provisions that are most relevant for women might be less salient for men. When we measure labor supply by annual earnings, we do not find statistically significant elasticities. The point estimate for the whole sample, 0.09, is positive and the corresponding 95%-confidence interval, -0.35 to $+0.52$, 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 both the extensive and intensive margin of labor supply. All the estimates in Table 5 are consistent with at least small incentive effects in the expected direction, while in four of the nine specifications we can reject the hypothesis of no incentive effects at the five percent level or better. It is important to keep in mind that our estimates are driven by responses to five discontinuities in the Social Security rules. If these discontinuities are particularly salient, then the behavioral response that we find could be greater than the behavioral response to benefit rules more generally. Moreover, our estimates are local treatment effects since they represent the responses of those individuals who are close to one or more of the discontinuities. It is possible that these individuals respond differently to labor supply incentives than the general population. Our estimates of the behavioral response to labor supply incentives may therefore not apply to other settings.

While we can reject the hypothesis that the link between earnings and future Social Security benefits is completely ignored in people's labor supply decisions, our results provide very limited guidance as to

[21] The Social Security benefits rules may also affect labor supply along the intensive margin through the earnings test, which reduces current benefits for individuals who have already taken up benefits but earn more than a certain threshold. However, reductions in current benefits due to the earnings test are compensated by increases in future benefits in a way that is roughly actuarially fair. Thus, labor supply responses to the earnings test indicate a lack of understanding of the tax-benefit link implicit in the earnings test. Friedberg (2000), Song and Manchester (2007), Haider and Loughran (2008), and Seif (2009) find evidence of labor supply responses to the earnings test, while Gruber and Orszag (2003) conclude that there is no robust evidence of labor supply responses by men and only suggestive evidence of labor supply responses by women.

Table 3a
Regressions of "smoothed" incentives on control variables.

Table reports: R^2 (top number) and root MSE (bottom number)	Sample for retirement regressions Dependent variable: smoothed ENTS			Sample for hours regressions Dependent variable: smoothed INTS			Sample for earnings regressions Dependent variable: smoothed INTS		
	Entire sample	Men	Women	Entire sample	Men	Women	Entire sample	Men	Women
1. No controls	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0455	0.0414	0.0209	0.0375	0.0416	0.0210	0.0386	0.0430	0.0216
2. Only basic demographics	0.2384	0.3678	0.6853	0.3306	0.4519	0.5458	0.3422	0.4570	0.5609
	0.0397	0.0330	0.0117	0.0307	0.0309	0.0142	0.0313	0.0317	0.0143
3. Baseline minus earnings history	0.8149	0.8006	0.8230	0.7668	0.8335	0.7815	0.7731	0.8394	0.7979
	0.0198	0.0188	0.0090	0.0184	0.0174	0.0101	0.0186	0.0176	0.0099
4. Baseline minus higher-order terms	0.9260	0.9082	0.8558	0.9178	0.9366	0.8373	0.9193	0.9366	0.8535
	0.0126	0.0128	0.0082	0.0110	0.0108	0.0088	0.0112	0.0111	0.0085
5. Baseline controls	0.9464	0.9380	0.8655	0.9488	0.9728	0.8458	0.9506	0.9721	0.8627
	0.0107	0.0106	0.0079	0.0087	0.0071	0.0086	0.0087	0.0074	0.0083
6. Baseline plus additional interactions	0.9743	0.9638	0.9558	0.9741	0.9847	0.9331	0.9753	0.9843	0.9401
	0.0075	0.0081	0.0046	0.0062	0.0054	0.0057	0.0062	0.0056	0.0055
7. Baseline plus further additional interactions	0.9766	0.9670	0.9587	0.9769	0.9866	0.9380	0.9778	0.9862	0.9444
	0.0072	0.0078	0.0045	0.0059	0.0051	0.0055	0.0059	0.0053	0.0054

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

how much of this link is perceived on average. To quantify the extent to which this link is perceived, we need to compare our estimated elasticities to labor elasticities from the literature that apply to the same population and that are identified off of incentive changes that are fully perceived. While we know of no estimates that satisfy these criteria exactly, labor supply elasticities are often estimated to be around 0.1 to 0.4 (Blundell and MaCurdy, 1999; Gruber and Saez, 2002; Kopczuk, 2005). Given this range and the confidence intervals on our estimates, our estimates are consistent both with a full perception of the tax-benefit link and with a very limited perception of this link.

4.5. Robustness

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

Rows 6 and 7 show that the baseline estimates are reasonably robust to making the control function substantially more flexible (the number of control variables almost doubles from the baseline specification to the specification in row 7). The estimates for the retirement regressions are extremely robust to making the control function more flexible. This is not surprising since estimates based on true discontinuities are not sensitive to the specification of the control

Table 3b
Regressions of true incentives on control variables.

Table reports: R^2 (top number) and root MSE (bottom number)	Sample for retirement regressions Dependent variable: ENTS			Sample for hours regressions Dependent variable: INTS			Sample for earnings regressions Dependent variable: INTS		
	Entire sample	Men	Women	Entire sample	Men	Women	Entire sample	Men	Women
1. No controls	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0495	0.0505	0.0467	0.0434	0.0459	0.0359	0.0450	0.0477	0.0376
2. Only basic demographics	0.0852	0.0721	0.1531	0.1376	0.0607	0.2704	0.1317	0.0570	0.2498
	0.0474	0.0487	0.0430	0.0403	0.0445	0.0307	0.0420	0.0463	0.0326
3. Baseline minus earnings history	0.4366	0.5150	0.4748	0.5743	0.5031	0.7271	0.5794	0.5209	0.7125
	0.0376	0.0357	0.0345	0.0288	0.0330	0.0192	0.0296	0.0336	0.0206
4. Baseline minus higher-order terms	0.5492	0.5918	0.5461	0.6515	0.5895	0.7716	0.6499	0.5977	0.7541
	0.0338	0.0330	0.0324	0.0262	0.0303	0.0178	0.0271	0.0311	0.0193
5. Baseline controls	0.5770	0.6177	0.5701	0.6788	0.6178	0.7818	0.6759	0.6224	0.7661
	0.0327	0.0320	0.0316	0.0251	0.0293	0.0174	0.0261	0.0301	0.0188
6. Baseline plus additional interactions	0.6296	0.6728	0.6083	0.7218	0.6699	0.8139	0.7179	0.6718	0.7989
	0.0308	0.0298	0.0305	0.0236	0.0275	0.0163	0.0245	0.0284	0.0176
7. Baseline plus further additional interactions	0.6443	0.6838	0.6287	0.7416	0.6944	0.8264	0.7361	0.6944	0.8096
	0.0304	0.0295	0.0299	0.0229	0.0266	0.0159	0.0239	0.0275	0.0173

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

Table 4
Effects with different sets of control variables, "smoothed" incentives.

Controls	Retirement regressions dependent variable: retirement			Hours regressions dependent variable: ln(Hours)			Earnings regressions dependent variable: ln(Earnings)		
	Entire sample	Men	Women	Entire sample	Men	Women	Entire sample	Men	Women
1. No controls	0.658***	0.965***	1.046***	0.353**	−0.326**	−2.651***	−6.117***	−8.461***	−12.834***
	(0.075)	(0.118)	(0.277)	(0.140)	(0.165)	(0.351)	(0.266)	(0.314)	(0.880)
2. Only basic demographics	−0.056	−0.207	0.076	1.700***	0.737***	−1.840***	−0.476***	−1.818***	−4.559***
	(0.088)	(0.150)	(0.446)	(0.162)	(0.211)	(0.430)	(0.176)	(0.273)	(0.678)
3. Baseline minus earnings history	−0.028	−0.146	0.246	1.104***	0.522	−0.731	−1.024***	−1.089**	−3.090***
	(0.167)	(0.234)	(0.553)	(0.237)	(0.328)	(0.510)	(0.310)	(0.450)	(0.899)
4. Baseline minus higher-order terms	0.386	0.141	0.816	0.060	−0.220	−0.203	−0.890*	−0.601	−1.139
	(0.242)	(0.320)	(0.577)	(0.381)	(0.540)	(0.556)	(0.537)	(0.776)	(0.928)
5. Baseline controls	0.272	−0.428	0.805	0.354	0.784	−0.354	−0.618	0.018	−0.588
	(0.284)	(0.382)	(0.595)	(0.457)	(0.800)	(0.520)	(0.677)	(1.153)	(0.970)
6. Baseline plus additional interactions	−0.030	−0.420	0.303	−0.383	−0.334	−0.828	−0.941	0.434	−0.976
	(0.395)	(0.485)	(0.990)	(0.560)	(0.908)	(0.758)	(0.877)	(1.522)	(1.264)
7. Baseline plus further additional interactions	0.007	−0.360	0.466	0.212	0.692	−0.378	−0.615	0.636	−0.543
	(0.407)	(0.503)	(0.957)	(0.558)	(0.903)	(0.735)	(0.894)	(1.587)	(1.345)

Notes: The table reports the coefficient and standard error on the "smoothed" INTS for ln(Earnings) and ln(Hours) regressions, and on "smoothed" ENTS for retirement regressions. Hours regressions are limited to individuals who report working at least 15 h per week in a typical week. Earnings regressions are limited to individuals with at least $2500 in earnings in the sample year. All dollars are 2003 dollars. Both hours and earnings regressions are limited to observations with predicted earning below the maximum taxable earnings for the OASDI tax. Standard errors, clustered by individual, are in parentheses. *indicates p-value<.10; **indicates p-value<.05; ***indicates p-value<.01. Appendix Table 2 contains detailed information on the exact set of control variables included in each row.

function (as long as it is continuous) and since the sharpest discontinuities occur in the extensive-margin incentives where there is no uncertainty about future own labor supply. The hours elasticity for the whole sample remains statistically significant in row 6 but becomes statistically insignificant in row 7. The hours elasticity for men, however, retains it statistical significance in both rows. The finding that the hours elasticity is somewhat more sensitive to the control function may be due to the fact that a very flexible control function will also absorb some of the non-linearities in incentives that have not fully converged to discontinuities.

Table 5
Baseline specification.

Dependent variable	Entire sample	Men	Women	p-value men=women
1. Retirement				
Coefficient on ENTS	−0.202**	−0.075	−0.323***	0.143
(Standard error)	(0.080)	(0.112)	(0.119)	
R^2	0.229	0.250	0.247	
N	13,902	7975	5927	
2. ln(Hours)				
Coefficient on INTS	0.415***	0.691***	−0.331	0.003***
(Standard error)	(0.158)	(0.176)	(0.309)	
R^2	0.280	0.245	0.303	
N	10,840	5891	4949	
3. ln(Earnings)				
Coefficient on INTS	0.090	0.021	0.382	0.303
(Standard error)	(0.220)	(0.264)	(0.428)	
R^2	0.514	0.517	0.526	
N	11,062	5984	5078	

Notes: Independent variable is INTS for ln(Earnings) and ln(Hours) regressions, ENTS for retirement regressions. Hours regressions are limited to individuals who report working at least 15 h per week in a typical week. Earnings regressions are limited to individuals with at least $2500 in earnings in the sample year. All dollars are 2003 dollars. Both hours and earnings regressions are limited to observations with predicted earning below the maximum taxable earnings for the OASDI tax. Standard errors, clustered by individual, are in parentheses. *indicates p-value<.10; **indicates p-value<. 05; ***indicates p-value<.01. Final column is a t-test of the equality of the coefficient on men and women. Appendix Table 2 contains detailed information on the exact variables included in the baseline set of controls.

In Table 7, we explore which discontinuities-in-the-limit contribute most to our estimates. Rather than use the variation from only a single provision of the Social Security benefit rules at the time, which yields very imprecise estimates, we examine the effect of "turning off" one provision at a time. We turn off single provisions by replacing each provision with its "smoothed" version as described in Section 2.3. It is important to keep in mind that there are many interactions between the provisions that create discontinuities-in-the-limit, and thus turning off a single provision will reduce the identifying power of other provisions. For example, turning off the 35-year rule modifies our estimates of the PIAs for individuals with an earnings history exceeding 35 years, which in turn causes the discontinuities from the rules on spousal and widow(er) benefits (which depend on PIA ratios) to appear in the wrong location. Despite this caveat, the magnitudes of the estimated incentive effects are relatively stable to turning off a single discontinuity-in-the-limit at the time, though we lose statistical significance in a number of cases. Turning off the 35-year rule has the largest impact on the size and significance of the estimated labor supply responses. Turning off either the provision on spousal benefits or the kinks in the AIME-PIA schedule has a smaller impact, while turning off either the vesting rule or the provision on widow(er) benefits has only a minimal impact on the estimated incentive effects. Overall, we conclude that interactions between the five provisions that generate discontinuities-in-the limit contribute importantly to our estimates, and that our estimates are not all driven by a single discontinuity-in-the limit.

Table 8 presents a final set of robustness tests. Panels A, B, and C show the robustness checks to the retirement regression, the hours regression, and the earnings regression, respectively. The first row of each panel reproduces the baseline regression. The second row shows that results are very robust to including the smoothed incentive as a control variable, which further confirms that the estimates are driven by the discontinuities-in-the-limit. In rows 3 and 4, we use alternative retirement definitions, based exclusively on earnings (row 3) or exclusively on self-reports (row 4) rather than our baseline definition, which combines

Table 6
Effects with different sets of control variables, true incentives.

Controls	Retirement regressions Dependent variable: retirement			Hours regressions Dependent variable: ln(Hours)			Earnings regressions Dependent variable: ln(Earnings)		
	Entire sample	Men	Women	Entire sample	Men	Women	Entire sample	Men	Women
1. No controls	−0.195***	−0.361***	−0.006	1.190***	0.693***	0.740***	1.238***	−0.697*	2.620***
	(0.066)	(0.093)	(0.100)	(0.125)	(0.144)	(0.256)	(0.302)	(0.367)	(0.534)
2. Only basic demographics	−0.127*	−0.075	−0.222**	1.031***	0.754***	0.419	−0.396***	−1.018***	−0.089
	(0.070)	(0.092)	(0.112)	(0.136)	(0.154)	(0.264)	(0.133)	(0.172)	(0.253)
3. Baseline minus earnings history	−0.072	−0.063	−0.192	0.625***	0.575***	0.189	0.278	−0.240	1.645***
	(0.075)	(0.102)	(0.127)	(0.141)	(0.159)	(0.312)	(0.198)	(0.238)	(0.394)
4. Baseline minus higher-order terms	−0.219***	−0.084	−0.361***	0.312**	0.559***	−0.441	0.058	−0.021	0.300
	(0.078)	(0.108)	(0.119)	(0.156)	(0.173)	(0.311)	(0.211)	(0.256)	(0.412)
5. Baseline controls	−0.202**	−0.075	−0.323***	0.415***	0.691***	−0.331	0.090	0.021	0.382
	(0.080)	(0.112)	(0.119)	(0.158)	(0.176)	(0.309)	(0.220)	(0.264)	(0.428)
6. Baseline plus additional interactions	−0.228***	−0.012	−0.361***	0.316**	0.606***	−0.381	0.054	0.041	0.242
	(0.085)	(0.121)	(0.128)	(0.160)	(0.177)	(0.328)	(0.241)	(0.290)	(0.460)
7. Baseline plus further additional interactions	−0.199**	−0.015	−0.321**	0.266	0.525***	−0.288	0.042	0.023	0.237
	(0.085)	(0.122)	(0.129)	(0.162)	(0.183)	(0.339)	(0.245)	(0.295)	(0.467)

Notes: The table reports the coefficient and standard error on the INTS for ln(Earnings) regressions, the INTS for ln(Hours) regressions, and the ENTS for retirement regressions. Hours regressions are limited to individuals who report working at least 15 h per week in a typical week. Earnings regressions are limited to individuals with at least $2500 in earnings in the sample year. All dollars are 2003 dollars. Both hours and earnings regressions are limited to observations with predicted earning below the maximum taxable earnings for the OASDI tax. Standard errors, clustered by individual, are in parentheses. *indicates p-value<.10; **indicates p-value<.05; ***indicates p-value<.01. Appendix Table 2 contains detailed information on the exact set of control variables included in each row.

information from both sources. We find that the results are somewhat sensitive to the choice of the retirement definition. For the retirement definition based exclusively on earnings, the labor supply response is no longer statistically significant along the retirement margin but retains significance along the hours margin. For the retirement definition based exclusively on self-reports, the labor supply response is no longer statistically significant along the hours margin and becomes only marginally significant along the retirement margin. 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 topcoding the ENTS at 0.10 rather than 0.50 in row 7). 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) contributes importantly to the magnitude of the baseline estimate of the hours regression for the whole sample. Topcoding the INTS for these observations at 0.10 decreases the point estimate by almost half and, as a result, this estimate is no longer statistically significant. The estimate in the subsample of men, however, remains significant.

5. Conclusion

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

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

Table 7
Sources of variation.

	Retirement regressions Dependent variable: retirement			Hours regressions Dependent variable: ln(Hours)		
	Entire sample	Men	Women	Entire sample	Men	Women
1. Baseline specification	−0.202**	−0.075	−0.323***	0.415***	0.691***	−0.331
	(0.080)	(0.112)	(0.119)	(0.158)	(0.176)	(0.309)
2. Incentives calculated without widow rules	−0.190***	−0.053	−0.221**	0.479**	0.951***	−0.365
	(0.072)	(0.125)	(0.091)	(0.196)	(0.241)	(0.302)
3. Incentives calculated without rules on spousal benefits	−0.172**	−0.101	−0.168*	0.158	0.708***	−0.554*
	(0.074)	(0.121)	(0.095)	(0.177)	(0.194)	(0.324)
4. Incentives calculated without 35-year rule	−0.153	−0.024	−0.201*	0.236	0.541	−0.258
	(0.109)	(0.124)	(0.116)	(0.243)	(0.389)	(0.303)
5. Incentives calculated without vesting rule	−0.194**	−0.111	−0.327**	0.414**	0.702***	−0.399
	(0.097)	(0.134)	(0.150)	(0.165)	(0.182)	(0.323)
6. Incentives calculated without kinks in the AIME-PIA schedule	−0.161*	−0.024	−0.249*	0.328**	0.612***	−0.221
	(0.089)	(0.131)	(0.133)	(0.166)	(0.182)	(0.321)

Notes: The table reports the coefficient and standard error on the INTS for ln(Earnings) regressions and the INTS for ln(Hours) regressions. Hours regressions are limited to individuals who report working at least 15 h per week in a typical week and with predicted earning below the maximum taxable earnings for the OASDI tax. Standard errors, clustered by individual, are in parentheses. *indicates p-value<.10; **indicates p-value<.05; ***indicates p-value<.01. The first row reproduces the baseline specification from Table 5. The remaining rows are identical to the baseline specification except that one of the Social Security benefit rules is turned off when calculating the incentive measure.

Table 8
Robustness checks.

	Entire sample	Men	Women
Panel A: retirement regressions (ENTS)			
1. Baseline	−0.202** (0.080)	−0.075 (0.112)	−0.323*** (0.119)
2. Smoothed retirement incentive controls	−0.236*** (0.082)	−0.040 (0.118)	−0.368*** (0.121)
3. Earnings-based retirement definition	−0.125 (0.080)	−0.046 (0.112)	−0.167 (0.121)
4. Retirement definition based on self-reports	−0.180* (0.096)	−0.325** (0.137)	−0.113 (0.148)
5. Probit of baseline regression	−0.192*** (0.074)	−0.115 (0.105)	−0.254** (0.100)
6. No windsorization of ENTS	−0.131*** (0.039)	−0.077 (0.057)	−0.171*** (0.064)
7. Windsorize ENTS at 0.10 (top 1.7%)	−0.261** (0.104)	−0.113 (0.153)	−0.365** (0.154)
Panel B: hours regressions (INTS)			
1. Baseline	0.401** (0.158)	0.691*** (0.176)	−0.331 (0.309)
2. Smoothed retirement incentive controls	0.396** (0.162)	0.682*** (0.177)	−0.307 (0.319)
3. Earnings-based retirement definition	0.381** (0.165)	0.716*** (0.183)	−0.375 (0.306)
4. Retirement definition based on self-reports	0.144 (0.151)	0.410** (0.162)	−0.450 (0.319)
5. Hours ≥10	0.325* (0.169)	0.686*** (0.182)	−0.602* (0.353)
6. Hours ≥20	0.292* (0.151)	0.521*** (0.170)	−0.249 (0.292)
7. Windsorize INTS at 0.10 (top 1.2%)	0.216 (0.149)	0.420** (0.164)	−0.252 (0.304)
Panel C: earnings regressions (INTS)			
1. Baseline	0.090 (0.219)	0.021 (0.264)	0.382 (0.428)
2. Smoothed retirement incentive controls	0.139 (0.227)	0.021 (0.276)	0.453 (0.429)
3. Earnings-based retirement definition	0.227 (0.235)	0.144 (0.279)	0.477 (0.463)
4. Retirement definition based on self-reports	−0.115 (0.227)	0.130 (0.278)	−0.374 (0.472)
5. Cut-off $5000	0.004 (0.202)	−0.057 (0.242)	0.251 (0.392)
6. Cut-off $1000	0.134 (0.239)	0.038 (0.288)	0.594 (0.467)
7. Windsorize INTS at 0.10 (top 1.2%)	0.127 (0.256)	−0.061 (0.309)	0.678 (0.497)

Notes: Independent variable is ENTS in Panel A and INTS in Panels B and C. All dollars are 2003 dollars. Standard errors, clustered by individual, are in parentheses. *indicates *p*-value<.10; **indicates *p*-value<.05; ***indicates *p*-value<.01. Regression controls are as in Table 5. Probit regressions report marginal effects. Windsorizing ENTS at 0.10 affects 1.7% of observations. Windsorizing INTS at 0.10 affects 1.2% of observations. Except where noted, regressions in Panel B are limited to individuals who work at least 15 h in a typical week. "Hours ≥ 10" regressions limit the sample to individuals who work at least 10 h in a typical week, likewise for "Hours ≥ 20" regressions. Except where noted, regressions in Panel C are limited to individuals with at least $2500 in earnings in the sample year. "Cut-off $5000" regressions limit the sample to individuals reporting at least $5000 of earnings in the year of observation, likewise for "cut-off $1000" regressions. Both hours and earnings regressions are limited to observations with predicted earning below the maximum taxable earnings for the OASDI tax. Retirement definitions are explained in Appendix A.

from increasing the transparency of the link between Social Security benefits and taxes may be smaller than is generally assumed.

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

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jpubeco.2009.07.006.

References

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

Blau, David M., 1997. Social Security and the labor supply of older married couples. Labour Economics 4 (4), 373–418.

Blinder, Alan S., Gordon, Roger H., Wise, Donald E., 1980. Reconsidering the work disincentive effects of social security. National Tax Journal 33 (4), 431–442.

Blundell, Richard, MaCurdy, Thomas, 1999. Labor supply: a review of alternative approaches. In: Ashenfelter, Orley, Card, David (Eds.), Handbook of Labor Economics, vol. 3. Elsevier, Amsterdam.

Boskin, Michael J., Kotlikoff, Laurence J., Puffert, Douglas J., Shoven, John B., 1987. Social Security: a financial appraisal across and within generations. National Tax Journal 40 (1), 19–34.

Brown, Jeffrey R., Liebman, Jeffrey B., Pollet, Joshua, 2002. Estimating life tables that reflect socioeconomic differences in mortality. In: Feldstein, Martin, Liebman, Jeffrey B. (Eds.), The Distributional Aspects of Social Security and Social Security Reform. InUniversity of Chicago Press, Chicago.

Browning, Edgar K., 1975. Labor supply distortions of Social Security. Southern Economic Journal 42 (2), 243–252.

Burkhauser, Richard V., Turner, John A., 1985. Is the Social Security payroll tax a tax? Public Finance Quarterly 13 (3), 253–267.

Burtless, Gary, 1986. Social Security, unanticipated benefit increases, and the timing of retirement. Review of Economic Studies 53 (5), 781–805.

Burtless, Gary, Moffitt, Robert, 1984. The effect of Social Security benefits on the labor supply of the aged. In: Aaron, Henry, Burtless, Gary (Eds.), Retirement and Economic Behavior. Brookings Institution, Washington, D.C.

Butrica, Barbara A., Johnson, Richard W., Smith, Karen E., Steuerle, C. Eugene, 2006. The implicit tax on work at older ages. National Tax Journal 59 (2), 211–234.

Chan, Sewin, Stevens, Ann Huff, 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, 2004. Retirement incentives and couples' retirement decisions. Topics in Economic Analysis and Policy 4 article 17.

Coile, Courtney, Gruber, Jonathan, 2007. Future Social Security entitlements and the retirement decision. Review of Economics and Statistics 89 (2), 234–246.

Coile, Courtney, Diamond, Peter, Gruber, Jonathan, Jousten, Alain, 2002. Delays in claiming Social Security benefits. Journal of Public Economics 84 (3), 357–385.

Cushing, Matthew J., 2005. Net marginal social security tax rates over the life cycle. National Tax Journal 58 (2), 227–245.

Diamond, Peter, Gruber, Jonathan, 1999. Social Security and retirement in the U.S. In: Gruber, Jonathan, Wise, David (Eds.), Social Security and Retirement around the World. University of Chicago Press, Chicago.

Feldstein, Martin S., Liebman, Jeffrey B., 2002. Social Security. In: Auerbach, Alan J., Feldstein, Martin (Eds.), Handbook of Public Economics, vol. 4. Elsevier, Amsterdam.

Feldstein, Martin S., Samwick, Andrew A., 1992. Social Security rules and marginal tax rates. National Tax Journal 45 (1), 1–22.

Fields, Gary S., Mitchell, Olivia S., 1984. Economic determinants of the optimal retirement age: an empirical investigation. Journal of Human Resources 19 (2), 245–262.

Friedberg, Leora, 2000. The labor supply effects of the Social Security earnings test. Review of Economics and Statistics 82 (1), 48–63.

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, Shoven, John B., Slavov, Sita Nataraj, 2009. Removing the disincentives in Social Security for long careers. In: Jeffrey, R., Brown, Liebman, Jeffrey, B., Wise, David A. (Eds.), Social Security Policy in a Changing Environment. University of Chicago Press, Chicago.
Gruber, Jonathan, Orszag, Peter, 2003. Does the Social Security earnings test affect labor supply and benefits receipt? National Tax Journal 56 (4), 755–773.
Gruber, Jonathan, Saez, Emmanuel, 2002. The elasticity of taxable income: evidence and implications. Journal of Public Economics 84 (1), 1–32.
Gustman, Alan L., Steinmeier, Thomas L., 1986. A structural retirement model. Econometrica 54 (3), 555–584.
Gustman, Alan L., Steinmeier, Thomas L., 2005a. The Social Security early retirement age in a structural model of retirement and wealth. Journal of Public Economics 89 (2–3), 441–463.
Gustman, Alan L., Steinmeier, Thomas L., 2005b. Imperfect knowledge of Social Security and pensions. Industrial Relations 44 (2), 373–397.
Gustman, Alan L., Mitchell, Olivia S., Samwick, Andrew A., Steinmeier, Thomas L., 1999. Pension and Social Security wealth in the health and retirement study. In: Smith, James, Willis, Robert (Eds.), Wealth, Work and Health, Innovations in Measurement in the Social Sciences. University of Michigan Press, Ann Arbor.
Hahn, Jinyong, Todd, Petra, Van der Klaauw, Wilbert, 2001. Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69 (1), 201–209.
Haider, Steven J., Loughran, David S., 2008. The effect of the Social Security earnings test on male labor supply: new evidence from survey and administrative data. Journal of Human Resources 43 (1), 57–87.
Hausman, Jerry A., Wise, David A., 1985. Social Security, health status, and retirement. In: Wise, David A. (Ed.), Pensions, Labor, and Individual Choice. University of Chicago Press, Chicago.
Kopczuk, Wojciech, 2005. Tax bases, tax rates and the elasticity of reported income. Journal of Public Economics 89 (11–12), 2093–2119.
Kotlikoff, Laurance J., 1996. Privatizing Social Security: how it works and why it matters. Tax Policy and the Economy 10, 1–32.
Krueger, Alan B., Pischke, Jörn-Steffen, 1992. The effect of Social Security on labor supply: a cohort analysis of the notch generation. Journal of Labor Economics 10 (4), 412–437.
Laitner, John, Silverman, Dan. 2008. "Consumption, Retirement and Social Security: Evaluating the Efficiency of Reform that Encourages Longer Careers," unpublished manuscript, University of Michigan.
Manoli, Dayanand, Mullen, Kathleen, and Wagner, Mathis. 2009. "Risk Aversion and Retirement Decisions: Using Policy Variation to Identify and Estimate a Structural Model of Retirement," Unpublished Manuscript, University of California — Los Angeles.
McCarty, Therese A., 1990. The effect of Social Security on married women's labor force participation. National Tax Journal 43 (1), 95–110.
RAND HRS Data, Version F, 2006. Produced by the RAND Center for the Study of Aging, with Funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA.
Rust, John, Phelan, Christopher, 1997. How Social Security and medicare affect retirement behavior in a world of incomplete markets. Econometrica 65 (4), 781–831.
Sabelhaus, John, 2007. What is the effective Social Security tax on additional years of work? National Tax Journal 60 (3), 491–506.
Samwick, Andrew A., 1998. New evidence on pensions, Social Security, and the timing of retirement. Journal of Public Economics 70 (2), 207–236.
Seif, David G. 2009. "Expected Mortality and the Social Security Earnings Test: Evidence from the 2000 Policy Change," Unpublished Manuscript, Harvard University.
Song, Jae G., Manchester, Joyce, 2007. New evidence on earnings and benefit claims following changes in the retirement earnings test in 2000. Journal of Public Economics 91 (3–4), 669–700.
Stock, James H., Wise, David A., 1990. Pensions, the option value of work, and retirement. Econometrica 58 (5), 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), 135–155.