Actuarial Fairness of Early Social Security Retirement Benefits[†]

Frank Heiland[‡]

Na Yin[♯]

September 23, 2011

Abstract

Reductions in Social Security retirement benefits for take-up before the Normal Retirement Age (NRA) are often assumed to be actuarially fair for average-mortality Americans. We investigate to what extent past, current, and (planned) future penalty schedules adhere to this principle. We show that, for a constant time discount rate, actuarial fairness requires the penalty schedule to be decreasing at a strictly increasing rate in take-up age. Assuming moderate levels of discounting, consistent with levels used by Social Security actuaries based on average real returns on long-term treasuries, the NRA 65 penalty schedule is found to be closest to actuarially fair for individuals who reached retirement age in the 1980s. For later cohorts subject to NRA 65, early retirement benefits have been increasingly less than actuarially fair despite a positive trend in interest rates that helped actuarially balance life expectancy gains. Following the 1983 Amendments, cohorts subject to NRA 66 and 67 face more balanced schedules. We also investigate additional reforms of the retirement ages and propose an amendment that would align future penalty schedules further with their actuarially fair form. We conclude with a discussion of the implications for optimal retirement behavior and the Social Security budget in light of recent benefit claiming patterns.

Keywords: Actuarial Fairness, Social Security Benefit Take-up, Retirement.

JEL classification: J26, H55.

[†]We are grateful to Cordelia Reimers, Hugo Benítez-Silva, and Sanders Korenman for helpful suggestions and discussions. We also received useful feedback during presentations at Baruch College, WEAI 2010, and EEA 2011. Edouard Chretien provided excellent research assistance. Heiland acknowledges financial support through PSC-CUNY Award 60121-40 41.
[‡] Corresponding Author: **Heiland:** CUNY Institute for Demographic Research, School of Public Affairs, CUNY Graduate Center (Economics Program), City University of New York, Baruch College, 1 Bernard Baruch Way, Box D-901, New York, NY 10010, phone: (646) 660-6868, fax: (646) 660-6871, e-mail: frank.heiland@baruch.cuny.edu
[♯] **Yin:** CUNY Institute for Demographic Research, School of Public Affairs, City University of New York, Baruch College, 1 Bernard Baruch Way, Box D-901, New York, NY 10010, phone: (646) 660-6824, fax: (646) 660-6871, e-mail: na.yin@baruch.cuny.edu

1 Introduction

The U.S. Social Security retirement program institutes sizable reductions in monthly benefits on individuals who elect to start receiving retirement benefits before reaching their Normal (or "Full") Retirement Age (NRA). The penalty for claiming benefits early depends only on a person's NRA, which varies by birth cohort, and on the number of months before the NRA benefits are claimed.[1] The maximal rate of reduction is for take-up at (exact) age 62, the Early Retirement Age (ERA) which has been in place since the early benefit take-up option became law (1956 for women and 1961 for men). For Americans turning 62 before the year 2000, that is those born before 1938, the NRA is (exact) age 65 and benefits are reduced linearly by 5/9 of 1% per month of earlier take-up. For subsequent cohorts, following the 1983 Amendments, the ERA has remained unchanged at 62 while the NRA has risen as shown in Figure 1. The NRA increased by 2 months per year across the 1938–1943 birth cohorts to (exact) age 66, and a similar increase is scheduled to be implemented for the 1954–1960 cohorts, resulting in an NRA of (exact) age 67 for Americans born in 1960 and thereafter. Cohorts after 1937 see benefits reduced by 5/9 of 1% per month of earlier take-up for the three year period before NRA and at 5/12 of 1% per month prior to that. Consequently, as illustrated in Figures 1 and 4, the penalty for collecting early benefits at a given age (62, 62 and 1 month, ..., 62 and 11 months, 63, ...) is increasing. For example, the maximal penalty has increased from 20% (=5/9*1%*36 months) for Americans reaching early retirement age before the year 2000 to 25% (=5/12*1%*12 months + 5/9*1%*36 months) for current cohorts of retirees (born 1943–1959) and it is scheduled to reach 30% (=5/12*1%*24 months + 5/9*1%*36 months) for those born in 1960 and thereafter.[2]

By allowing Americans to elect reduced benefits before their NRA, and as early as age 62, Social Security grants considerable financial flexibility in the timing of the first receipt of retirement benefits.

[1] Social Security covers the vast majority of individuals. For example, in 2010, 90% of Americans age 60 to 64 were considered fully insured (SSA-S, 2010, Table 4.C5).

[2] This paper does not consider spousal and widower benefits. The complexities introduced by those considerations are beyond the scope of the present analysis. See Myers and Schobel (1990), Gustman and Steinmeier (1991), Coile et al. (2002), Votruba (2003), Munnell and Soto (2005), and Sun and Webb (2009) for a discussion. By ignoring spousal and widow(er) benefits we are not taking into account the fact that approximately 6% of the individuals who receive some type of Old Age, Survivors, or Disability Insurance (OASDI) benefits receive them as spouses of entitled retirees. This percentage comes from the 2001 Public-Use Master Beneficiary Microdata provided by the Social Security Administration, a 1% random sample of all beneficiaries in December of 2001.

When deciding whether to retire at 62 (or any age before NRA), individuals choose between a discounted expected stream of smaller (monthly) payments from Social Security that starts immediately and a stream of larger benefit checks that starts at a future date. This trade-off is fundamentally influenced by the size of the age-related reductions (henceforth: "the penalty schedule") and individuals' (expected) mortality, in addition to individuals' preferences for time (discounting) and risk. When benefit collection begins and how long benefits are received directly influences the outlays (and to a lesser extent the contributions) of Social Security. One principle to consider in the design of the penalty structure is the actuarial fairness of individuals' lifetime benefits, a concept familiar from the pension and insurance literature. A penalty structure is actuarially fair for an eligible person with given (expected) mortality if the set of age-related reductions is such that the individual's expected discounted value of lifetime Social Security benefits (in real terms) does not vary by age at take-up (see Burkhauser, 1980; Myers and Schobel, 1990; Gruber and Wise, 2005). Or, as Gruber and Wise (2005, p.6) put it, *...if benefits are taken at age 64 instead of 65, they are reduced just enough to offset the receipt of benefits for one additional year.* Some authors use a slightly different terminology to denote the same concept.[3] The application of the actuarial fairness principle here is to be distinguished from contexts that use the same or similar terminology in the pension literature comparing benefits and contributions.[4]

The retirement literature has taken Social Security's age-related reductions for take-up before NRA as "approximately actuarially fair" for Americans with average mortality.[5] While there are no recent official statements or publications by the Social Security Administration on the actuarial properties of the penalty schedules to our knowledge, the actuaries at Social Security have repeatedly stated in the

[3]For example, Queisser and Whitehouse (2006) and Börsch-Supan (2000) use the term actuarial neutrality, Duggan and Soares (2002) use "actuarial equivalence," and Crawford and Lilien (1981) refer to "marginal fairness."

[4]With respect to the balance between contributions made and benefits received, the "actuarial fairness" has been used to describe pension systems where payments into the system equal benefits received for the average person (see e.g., Queisser and Whitehouse, 2006; Belloni and Maccheroni, 2006). As part of their solvency projections the Social Security Trustees use "actuarial balance" in reference to the difference between the present value of total tax income for a period (75 years in their longest projections) and the present value of the total cost (outlays) for the period, each divided by the present value of taxable payroll for all years in the period.

[5]See, for example, Crawford and Lilien (1981, p.515), Benítez-Silva and Heiland (2007, p.531), and Sun and Webb (2009, p.3). Gruber and Wise (2005, p.6, Footnote 2) write in a recent working paper: *Under current law, benefits in the United States are actuarially fair between 62 and 65, but are increased less than actuarially if the receipt of benefits is delayed beyond age 65, thus providing an incentive to leave the labor force at 65. Benefits will eventually become actuarially fair after age 65 as well.*

past that the NRA 65 schedule, affecting beneficiaries born before 1938, is (approximately) actuarially fair.[6] Robert Myers, the chief actuary of Social Security between 1947 and 1970, and his co-author projected in a 1990 paper that the reductions in the law "are reasonably close to the theoretically correct values" (the actuarially fair ones) for workers retiring in 1990 and in 2030 (see Myers and Schobel, 1990, p.295). Given the significant rise in average remaining life spans at age 62 and 80 between 1960 and 1999 (see Figure 2), four decades of consecutive birth cohorts facing identical penalties for early retirement benefit take-up, differences in how close to actuarially fair Social Security's early retirement penalties have been for average Americans are expected across cohorts.

There is little systematic research on the properties of the theoretically fair penalty schedule and few studies have analyzed how close Social Security's schedules are to being actuarially fair.[7] Duggan and Soares (2002) examine the actuarial fairness ("equivalence" in their terminology) of lifetime benefits across gender and earnings among U.S. beneficiaries born between 1900 and 1933 (NRA 65). Comparing actual ("statutory") and actuarially fair benefits, using within-sample mortality estimates and assuming a constant time discount rate of 3% per year, they find that men who claimed early (especially those with low earnings) enjoyed an actuarial premium. Women who claimed early, on the other hand, tend to face less than actuarially fair benefits, owing to their longer life expectancy. The paper also illustrates how higher discount rates require larger actuarial adjustments to be fair and it discusses the consequences of a higher NRA. Queisser and Whitehouse (2006) examine the actuarial fairness ("neutrality") of the *average* annual reductions for early and delayed benefit take-up across OECD countries. Assuming a 2% discount rate actuarially fair penalty schedules are calculated for a hypothetical pension system with flexible eligibility ages. The main result pertaining to U.S. mortality data (from 2002) is an actuarially fair *average* reduction for early take-up of 8.2% per year for NRA 67, exceeding the average statutory reduction of 6% (= 30% / 5 years) per year.

Duggan and Soares (2002) study beneficiaries born in 1900-1933, cohorts that faced the same penalty schedule (NRA 65), and their primary focus is on the deviations from fairness associated with

[6]Professor Reimers shared details to that effect with us from correspondence that she had with the actuaries during the 1990s.

[7]A handful of earlier studies exist that focus on one or two (groups of) cohorts (Blinder et al., 1980; Myers and Schobel, 1990; Mirer, 1998; Liu and Rettenmaier, 2006).

income and gender. Queisser and Whitehouse (2006) look at the average annual penalty not the actual penalty schedule(s) and their U.S. results only pertain to the fairness of the NRA 67 schedule (and only based on 2002 period mortality). Both studies abstract from the important issues of the shape of the penalty schedules and its evolution across cohorts as age-specific mortality rates decline. The present paper provides a more comprehensive, cohort-based analysis of the actuarial properties of Social Security's early retirement penalties. We conduct a formal analysis of the properties of actuarially fair penalty schedules and quantify how close the statutory penalties are to being actuarially fair for average mortality individuals in past, current, and future cohorts of U.S. retirees. The investigation considers various time discounting scenarios and analyzes the sensitivity of the conclusions to assumptions regarding cohort mortality. Moreover, we examine the consequences of likely retirement age reform scenarios (NRA of 69, ERA of 64) on the actuarial fairness of early retirement benefits and propose changes to the penalty formula that would align future penalty schedules more closely with an actuarially fair form.

Understanding the extent to which early benefits are actuarially fair for average Americans likely has important budgetary implications. Americans increasingly claim benefits before they reach their NRA and for some of the cohorts investigated here, more than half claim as soon as they reach age 62 (see Table 3 and Figure 13). At the same time, mortality rates past age 60 have fallen substantially over the past six decades, resulting in a rise in life expectancy conditional on reaching age 62 (80) from 78 (87) years in 1960 to 82 (88 and 1/2) years in 2000 (see Figure 2 and Figure 3). If benefits are (approximately) actuarially fair for individuals with average mortality in their cohorts, then the age at which they start collecting benefits would have little effect on the Social Security finances from a lifetime budgetary perspective.[8] From a cash-flow (pay-as-you-go) perspective, on the other hand, Social Security always prefers delay in benefit take-up. If, however, the penalty schedule results in early benefits that are discounted too much, then the fact that beneficiaries tend to claim benefits early may lower Social Security's expected long-term payout burden.

[8]The program's finances could still be negatively affected as a result of the loss in payroll taxes that the early claimers would have paid into the system if they had continued to work and delayed benefit take-up. Individuals with 35 years of earnings under the program can stop working without reducing their pension wealth much but the program loses any years where contributions were made.

Our analysis also relates to the broader retirement literature.[9] While this literature has paid relatively little attention to the incentives provided by the penalty schedule in conjunction with the secular decline in mortality, it has studied a number of factors that influence the timing of retirement and benefit take-up including variables related to life expectancy. For example, using data from the Social Security Master Beneficiary Record and the Health and Retirement Study (HRS) data, Coile et al. (2002) find that men with higher life expectancy delay benefit take-up. Analyzing subjective mortality data from the HRS, Hurd et al. (2002) find that the subjective survival forecasts predict actual mortality as well as benefit take-up.[10] The present study may benefit Americans trying to make decisions regarding the optimal timing of benefit take-up. The age-related reductions are common to all beneficiaries from the same birth cohort.[11] Consequently they provide incentives that differ across individuals depending on their (expected) mortality. Knowing how the penalty schedule compares actuarially to the fair schedule for cohorts that face different (predicted) average mortality profiles provides information that can be useful in the determination of the optimal timing of benefit take-up by individuals. If the benefit schedule is approximately actuarially fair for the average mortality person in a given cohort, we can make predictions regarding the optimal take-up age of individuals with different characteristics facing the same benefit schedule. In particular, individuals facing below (above) average life expectancy should claim benefits early (late), holding other factors such as time and risk preferences constant.

Section 2 provides a brief overview of the rules governing Social Security retirement benefits with a focus on the reductions for early benefit claiming. Section 3 defines the concept of actuarial fairness and develops a simple framework to analyze it formally. Section 4 discusses important properties of actuarially fair penalty schedules for constant time discounting and assesses the degree of actuarial fairness of past, current, and future (planned and counterfactual) penalty schedules across cohorts using

[9] For a survey of the broader literature see Lumsdaine and Mitchell (1999). Hurd (1990), Lumsdaine (1995), and Ruhm (1996) provide good discussions of the earlier literature.

[10] Waldron (2004) estimates the correlation between mortality risk and retirement timing and finds that the majority of early retirees have higher mortality risk than age 65 retirees. Benítez-Silva and Heiland (2008) estimate joint models of benefit claiming and retirement after age 62 from HRS data and find some evidence that higher subjective survival forecasts predict a slower exit from the labor force.

[11] Anti-discrimination laws may make a system where the penalty schedule is gender- or race-specific impossible. Myers and Schobel (1990, p.298) note that *after the courts found various differences in the treatment of men and women under U.S. Social Security law to be unconstitutional, Congress prospectively removed virtually all such differences.*

different discount and mortality scenarios. Section 5 concludes with a discussion of the implications of the main findings for optimal retirement behavior and the Social Security budget.

2 Social Security Rules for Early Retirement

2.1 Benefit Calculation

Individuals aged 62 or older who had earned income subject to the Social Security payroll tax for at least 10 years since 1951 are eligible for retirement benefits under the Social Security Old Age (SSOA) benefits program. Earnings are subject to the tax up to an income maximum that is updated annually according to increases in the average wage.[12] To determine the monthly benefit amount (MBA), the Social Security Administration calculates the Primary Insurance Amount (PIA) of a worker as a convex piece-wise linear function of the worker's average earnings subject to Social Security taxes taken over her 35 years of highest earnings.

2.2 Early Benefit Take-up

Social Security has complex rules regarding the up-take of benefits. For our purpose the most important aspect is the penalty rate (PR), which measures the (permanent) reduction in benefits that individuals face if they claim benefits before reaching their NRA. In their publications Social Security uses the term Reduction Factor (RF) which is the amount of benefits received expressed as a fraction of PIA, that is RF equals 1-PR. If the benefits are claimed at NRA, the MBA equals the PIA. If an individual decides to claim benefits before the NRA and exits the labor force or stays below the limits prescribed by the Social Security Earnings Test as explained below, her MBA is reduced by the penalty rate.

Under the current rules of Social Security's Old Age Program, for a person with NRA of age 66 (those born between 1943 and 1954), the monthly benefit amount received upon claiming benefits

[12]Six percent of the 162.5 million workers with Social Security taxable earnings in 2008 had earnings at or above the maximum amount (SSA-S, 2010, Table 4.B1).

before the NRA depends on the age (in months) at take-up in the following way:

$$MBA_t = \begin{cases} (0.75 + 0.05 * \frac{1}{12} * (\text{Months not claimed after reaching 62})) * PIA \\ \qquad \text{if claimed before or at age 63;} \\ (0.80 + 0.05 * \frac{1}{9} * (\text{Months not claimed after reaching 63}))*PIA \\ \qquad \text{if claimed between age 63 and 66,} \end{cases}$$

where MBA_t represents the monthly benefit amount before the NRA (SSA-S, 2005, p.18). Assuming the typical case where individuals receive all monthly benefit checks, claiming early leads to a permanent reduction in a person's MBA_t.[13]

Underlying the calculation of the MBA is a reduction of benefits by 5/9 of 1% for each month benefits are claimed in the three years prior to NRA, plus a reduction of 5/12 of 1% for each month benefits are claimed before that. Thus at the current NRA of 66 the maximum reduction in benefits is 25% (PR=0.25) while it is scheduled to reach 30% (PR=0.30) for those born in 1960 and thereafter as the NRA increases to 67 over the next few years (see SSA-S, 2005, p.18). Cohorts born prior to 1938 faced an NRA of 65 and a constant marginal penalty of 5/9 of 1% per month for the entire three year period prior to age 65 for a maximum PR at age 62 of 20%. The 5/12 of 1% monthly reduction came as part of the Social Security Amendments of 1983 to increase the NRA.

According to Robert Myers, the chief actuary of the Social Security Administration from 1949 to 1970, the 5/9 of 1% per month rate of reduction introduced in legislation from 1956 was determined using linear interpolation after establishing a 20% maximum penalty at 62 at that time (see Myers and Schobel, 1990, p.295). During the deliberations for the 1983 Amendments, Congress apparently also considered a proposal to gradually increase the ERA in tandem with the NRA by (ultimately) two years to 64. This would have shifted the penalty schedule without changing its curvature. Sliding the three year period with the higher monthly penalty rate of 5/9 of 1% and assessing a rate of 5/12 of 1% before that may have been a compromise to keep the penalty percentages at the focal ages, namely 62 and 65, at relatively round numbers (25% and 30% at 62, 20% and 25% at 63 for NRA 66 and 67, respectively)

[13] The permanent reduction has been the norm. Benítez-Silva and Heiland (2007, p.537, Figure 1) report that in 2004, 6.6% of males and 5.8% of females had any benefits withheld due to earnings during their first year of as early beneficiaries.

when it became clear that ERA 64 was infeasible.[14]

Figure 1 shows the NRA and the penalty rate at age 62 for birth cohorts 1908-2032. It is important to note that the penalty rate only depends on how many months before NRA a person claims and her birth cohort (which determines her NRA). This information is summarized in the penalty rate schedule. Figure 4 charts the penalty schedules for NRA 65 (before 1999), NRA 66 (current) and NRA 67 (future according to 1983 Amendments). The figures illustrate how the penalty schedule has changed across cohorts. All retirees born before 1938 faced a penalty schedule that is linear in take-up age. Cohorts born after 1937 are facing a penalty schedule that is piece-wise linear with a smaller marginal penalty between ERA and NRA - 3 Years (5/12 of 1% per month) than between NRA - 3 Years and NRA (5/9 of 1% per month).

While the focus of this paper is on the actuarial fairness of the age-related adjustments before NRA, it is important to note that additional adjustments exist during the post-NRA period. Workers claiming benefits after the NRA earn the delayed retirement credit (DRC). For those born in 1943 or later it is 2/3 of 1% for each month up to age 70. For those born before 1943 it ranges from 11/24 to 5/8 of 1% per month, depending on their birth year. Observers generally consider the credits to be too small to be actuarially fair for retirees born before 1943 (see e.g., Myers and Schobel, 1990). The 8% increase per year of delay past NRA for those born in 1943 and thereafter, on the other hand, has been described as approximately fair by Myers and Schobel (1990), based on cohort mortality projected for 2030, which was confirmed in more recent analysis by Queisser and Whitehouse (2006).

2.3 Earnings Test

The rules governing early benefits allow for the adjustment of the penalty rate as a result of benefit withholding due to earnings above the Social Security Earnings Test limits (see Benítez-Silva and Heiland 2007 and 2008).[15] The Earnings Test determines the maximum earnings that do not result in a benefit reduction for individuals who have claimed retirement benefits before the NRA. Individuals

[14]See the discussion in Myers and Schobel (1990) on the preference of Congress for “round” numbers.

[15]Until its abolition in 2000 there was also a post NRA Earnings Test.

who continue (including re-enter) employment after claiming benefits early and who earn less than the earnings limit receive benefits at the reduced rate. Those with earnings above the limit, on the other hand, may not receive all monthly checks from Social Security, as some benefits will be withheld. The benefits that are banked this way will be redistributed through greater monthly benefits after NRA when the penalty rate is reset (decreased).[16]

Early Social Security benefits are withheld for earnings above the limit at a rate of 50% for beneficiaries between age 62 and the January of the year in which they reach the NRA, and 33% from January of that year until the month they reach the NRA (SSA-S, 2005, p.19; SSA-S, 2005, Table 2.A18). In the latter period, the earnings limit is higher, $37,680, compared with $14,160 for the earlier period as of 2009 (SSA-S, 2009, Table 2.A29).

It is important to emphasize that the (downward) adjustment of the penalty rate is automatic and becomes effective only after a person reaches NRA. Any benefits received prior to NRA will be at the initial reduced level. This leads to a peculiarity regarding the actuarial fairness of the age-related reduction (see Benítez-Silva and Heiland 2007 and 2008): For a person for whom the penalty schedule was originally actuarially fair who collects early benefits after having (some) benefits withheld, the penalty rate is no longer actuarially fair until the NRA.[17]

3 Assessing Actuarial Fairness

The mechanism of penalizing early retirement benefit take-up described in the previous section is at the heart of Social Security's early retirement rules and has been well-documented.[18] However, as we discuss in the introduction, there is relatively little research on the actuarial properties of the penalty schedules used by Social Security and detailed analyses of how close to actuarially fair the (changing)

[16]Benítez-Silva and Heiland (2007) provide additional detail on the rules governing benefit withholding and a numeric example of the streams of income resulting from these incentives. There are also other strategies for early claimers to reduce or potentially entirely undo the penalty effective at NRA (see Benítez-Silva and Heiland, 2007; Munnell et al., 2009).

[17]The related prior research has primarily focused on the taxation aspect of the Earnings Test. See Vroman (1985), Burtless and Moffitt (1985), Honig and Reimers (1989), Leonesio (1990), Reimers and Honig (1993), Reimers and Honig (1996), Friedberg (1998), Baker and Benjamin (1999), Friedberg (2000), and Votruba (2003).

[18]Leonesio (1990), Myers and Schobel (1990), Gustman and Steinmeier (1991), Myers (1993, p.52), Gruber and Orszag (2003), and Benítez-Silva and Heiland (2007, 2008) discuss the rules in some detail.

penalty schedule is for average mortality individuals who reach eligibility age across a wide range of birth cohorts are lacking.

3.1 Definition

Conducting a formal analysis of the actuarial properties of the penalty schedules requires that we operationalize the "individual of average mortality." We will consider the average or representative person in a cohort utilizing the following two types of mortality information: (1) the average life expectancy of the cohort of interest (conditional on reaching age 62) and (2) the average age-specific mortality risk profile (after age 62). Two (related) definitions of actuarially fair penalty schedules ensue:

Actuarial Fairness Definition 1 (Average Life Expectancy)
For a person who lives exactly to the age that coincides with the average life expectancy in the population (conditional on reaching ERA) the early retirement penalty schedule is actuarially fair if it equalizes that person's expected discounted value of lifetime benefits across take-up ages ERA, ERA+1, ..., NRA.

Actuarial Fairness Definition 2 (Average Age-Specific Mortality Risk)
For a person with average age-specific mortality after age ERA, the early retirement penalty schedule is actuarially fair if it equalizes that person's expected discounted value of lifetime benefits across take-up ages ERA, ERA+1, ..., NRA.

Definition 1 emerges as a special case of Definition 2 where the probability to survive to the next period is assumed to be 1 until age $NRA + D - 1$ and 0 thereafter, where D represents the number of years lived after reaching NRA. Definition 1 uses a single parameter, life expectancy, to summarize mortality in a population. Since average life expectancy conditional on reaching age 62 has always exceeded the NRA (see Figure 2), Definition 1 assumes that a person who reaches age 62 is certain to live past their NRA. It is easy to see how the penalty that preserves actuarial fairness as defined by Definition 1 is linked to the greater number of benefit periods that early claimers enjoy. Definition 2 considers an individual whose period-to-period survival before and after the NRA is uncertain. It takes a more accurate life course risk perspective and links average age-specific survival risk profiles to the age-specific benefit reduction schedule.

3.2 Present Value Method

Definitions 1 and 2 can be expressed formally using (expected) present values of lifetime benefit streams.[19] Discounting all present values to ERA, the condition for actuarial fairness of a penalty schedule for an individual with given age at death, $NRA+D$ (Definition 1), can be formally expressed as follows:

Life Expectancy Perspective

$\sum_{a=ERA}^{NRA+D}(1-r(ERA))\cdot PIA = \sum_{a=ERA+1}^{NRA+D}(1-r(ERA+1))\cdot PIA = ... = \sum_{a=NRA}^{NRA+D}(1-r(NRA))\cdot PIA$, where $r(c)$ denotes the fraction by which a person's benefits are reduced relative to their full benefits at NRA when claiming at age c. The actuarial properties of the early penalty schedule ($c \in ERA,...,NRA$) are our main interest. We note that $1-r(c)$ represents the fraction of benefits paid by claiming age (the "benefit rate schedule"). Notice that since there is no penalty at NRA it is true that $r(NRA)=0$ and $1-r(NRA)=1$. Since a represents exact age in years, the penalty schedule, $r(c)$, is accurate for take-up the month the person reaches age ERA, ERA+1,..., NRA.

Assuming that future benefit streams are discounted at an annual rate δ, the condition for actuarially fair benefits becomes:

$\sum_{a=ERA}^{NRA+D}(1-r(ERA))\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot PIA = \sum_{a=ERA+1}^{NRA+D}(1-r(ERA+1))\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot PIA = ... = \sum_{a=NRA}^{NRA+D}(1-r(NRA))\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot PIA$.

To formally express Definition 2 we incorporate the uncertainty of survival between ages as follows:

Life Course Perspective

$\sum_{a=ERA}^{NRA+D}\prod_{i=ERA}^{a}S(i)\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot(1-r(ERA))\cdot PIA = \sum_{a=ERA+1}^{NRA+D}\prod_{i=ERA}^{a}S(i)\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot(1-r(ERA+1))\cdot PIA = ... = \sum_{a=NRA}^{NRA+D}\prod_{i=ERA}^{a}S(i)\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot(1-r(NRA))\cdot PIA$

[19] The idea that individuals' benefit claiming decisions are affected by the (expected) asset value of their Social Security pension wealth, as assumed in the present value approach, has good empirical support (see e.g., Burkhauser, 1980).

where $S(a)$ is the probability of being alive at age a conditional on survival to age $a-1$. Since we focus on the population age 62 and above, we set $S(ERA) = 1$.

The definition above assumes a risk-neutral individual. This is appropriate from the perspective of the Social Security actuaries since their focus is on aggregate (average and realized) outcomes (see Myers and Schobel, 1990). In contrast, there is the perspective of individuals trying to solve their retirement problem, comparing expected lifetime benefit streams given their particular constraints (e.g., mortality) and preferences for time and risk. To capture the risk preference characteristics of Americans solving their retirement problem, we would want to assume risk-averse preferences, recognizing that individuals will prefer a certain amount x over a lottery with expected value x. The definition of the fair schedule for this more general life course case can be found in Appendix A.

The discount rate is an important determinant of the fair penalty schedule. The assumption of a constant discount rate is for expositional convenience and we also consider cases of variable discounting in the analysis below. We note that the discount rate appropriate in the context of the actuarial calculations by Social Security is unlikely to coincide with a subjective time preference rate that is realistic for average Americans trying to solve their retirement problem (see Duggan and Soares, 2002). Individuals are likely to discount the future more heavily than actuaries would on behalf of Social Security, to account for risks related to health (e.g., mortality) and economic wellbeing (e.g., borrowing constraints).

4 Results

Our analysis of the actuarial properties of the penalty schedules proceeds in three main steps. First, we formally derive and discuss the properties of the schedules implied by each definition. Second, we compare schedules obtained from calibrations of the life expectancy perspective (Definition 1) to the actual schedule for the NRA 65 schedule. Third, we implement the life course perspective (Definition 2) and assess the actuarial fairness of past, current, and future planned and counterfactual penalty schedules across cohorts by comparing present values of benefit streams by age at benefit take-up.

The analysis abstracts from differences in the real pension wealth across cohorts to facilitate the comparison of the actuarial properties of the penalty schedule across cohorts. Over time average PIAs have increased somewhat in real terms, primarily as a result of real wage gains, rising labor force participation (especially among women), and program expansion.[20] We note that in the U.S. system, after age 60 the PIA is indexed to the price level using a cost-of-living-adjustment. By assuming the same (real) PIA across cohorts, we also abstract from potential increases in the PIA after age 62 that can occur when a person's earnings are sufficiently high to replace a prior year of lower earnings.

We focus on take-up at exact ages 62, 63, ..., NRA. Abstracting from the interim months simplifies the formal exposition. The results for take-up during interim months can easily be approximated by interpolation. We also note that for an exact empirical assessment of take-up at interim months age-specific mortality data are needed which are typically not available.

The analysis employs publicly available U.S. mortality data from the Human Mortality Database (http://www.mortality.org/). For the analysis of the life expectancy case, we use estimates of average longevity conditional on survival to age 62 from Period Life Table data (see Figure 2). Cohort annual age-specific mortality data past age 60 are used to estimate the average age-specific survival probabilities in the life course analysis. Age-specific cohort mortality rates after 2005 are not available. As we discuss in more detail below, we impute missing values using data from adjacent cohorts and examine the sensitivity of the results using predicted cohort mortality from an auxiliary regression.

Social Security uses data on real risk-free interest rates to proxy for the discount rate in all its actuarial analyses, including the long-term solvency estimates for the Social Security Trust Fund. Similarly, we construct a real interest rate time series based on data on average annual yields of long-term constant coupon U.S. treasuries available from the Federal Reserve.[21]

[20]For example, Benítez-Silva and Yin (2009, p.7, Table 3) report that the average monthly benefits (in constant 2005 dollars) among retirees of age 62 increased from $1,066 in 1994 to $1,135 in 2004.

[21]Our real interest rate time series is based on data on average annual yields of long-term U.S. treasuries available online at www.federalreserve.gov (deflated using the CPI-U). Comparing our series to one provided by the Actuarial Office of Social Security we noticed that prior to 1999 they report only 5-year averages resulting in a significantly less volatile series. Social Security reports an average annual real interest rate of 2.8% based on data from the 1968-2007 period, closely matching our average of 2.76% based on 1954-2010 data.

4.1 Properties of a Fair Schedule

From the present value expressions stated above we derive several key properties of the actuarially fair penalty schedule. The following propositions are stated for the case of NRA 65.

Proposition 1 In the baseline case with discount rate δ (discount factor of $\beta = \frac{1}{1+\delta}$) the actuarially fair penalty schedule is $r(c) = \frac{1-\beta^{NRA-c}}{1-\beta^{NRA+D-c}}$. For example, given an NRA of 65 the penalties at ages 62, 63, and 64 are $\frac{1-\beta^3}{1-\beta^{D+3}}$, $\frac{1-\beta^2}{1-\beta^{D+2}}$, and $\frac{1-\beta}{1-\beta^{D+1}}$. No discounting emerges as a special limiting case where $\delta \to 0$ ($\beta = \frac{1}{1+\delta} \to 1$): In the baseline case without discounting (discount factor of 0%) the actuarially fair penalties at ages 62, 63, and 64 are $\frac{3}{D+3}$, $\frac{2}{D+2}$, and $\frac{1}{D+1}$.

We note that Proposition 1 suggests that the actuarially fair penalty schedule is strictly concave in age at benefit initiation before NRA (that is the fair early benefit schedule is strictly convex). This result is intuitive since earlier benefit initiation means that smaller checks (reduced benefits) are collected for a longer period of time.The longer the period when reduced benefits are collected, the smaller the penalty needs to be to make the person indifferent between claiming early and waiting until age of full benefits. By the same reasoning it is also clear that the penalty schedule shifts down (early claiming is less penalized) as life expectancy (D) increases. An increase in the number of periods after the NRA also results in a penalty schedule that is less concave (or less curved) in claiming age. Lastly we note the importance of the time discount rate, δ. The reduction schedule is lower for a higher discount rate (lower discount factor, β). Intuitively, to make individuals that value future benefits less indifferent between initiating a benefit stream today or tomorrow requires a greater penalty for early take-up. All else equal, a higher discount rate will require a more concave (or more curved) penalty schedule.

Proposition 2 In the life course case with risk-neutral individuals, discount rate δ (discount factor of $\beta = \frac{1}{1+\delta}$) and variable survival rate $S(a)$, the actuarially fair penalty at (early) claiming age c is $r(c) = \frac{\sum_{a=c}^{NRA} \delta^{a-ERA} \prod_{i=ERA}^{a} S(i)}{\sum_{a=c}^{NRA} \delta^{a-ERA} \prod_{i=ERA}^{a} S(i) + \sum_{a=NRA+1}^{NRA+D} \delta^{a-ERA} \prod_{i=ERA}^{a} S(i)}$. For example, given an NRA of 65 the fair penalties at ages 62, 63, and 64 are $\frac{1+\delta\cdot S(63)+\delta^2\cdot S(63)S(64)}{1+\delta\cdot S(63)+\delta^2\cdot S(63)S(64)+A}$, $\frac{\delta\cdot S(63)+\delta^2\cdot S(63)\cdot S(64)}{\delta\cdot S(63)+\delta^2\cdot S(63)\cdot S(64)+A}$, and $\frac{\delta^2\cdot S(63)\cdot S(64)}{\delta^2\cdot S(63)\cdot S(64)+A}$, where $A \equiv \sum_{a=65}^{65+D} \delta^{a-62} \prod_{i=62}^{a} S(i)$. (Recall that $S(ERA) = S(62) = 1$.) When the period survival transition probabilities are constant, $S(a) = \bar{s}$, the actuarially fair penalties at ages 62, 63, and 64 become $\frac{1-(\bar{s}\beta)^3}{1-(\bar{s}\beta)^{D+3}}$, $\frac{1-(\bar{s}\beta)^2}{1-(\bar{s}\beta)^{D+2}}$, and $\frac{1-\bar{s}\beta}{1-(\bar{s}\beta)^{D+1}}$. No discounting emerges as a special case where $\beta = 1$. The fair penalty schedule in this case is: $\frac{1-\bar{s}^3}{1-\bar{s}^{D+3}}$, $\frac{1-\bar{s}^2}{1-\bar{s}^{D+2}}$, and $\frac{1-\bar{s}}{1-\bar{s}^{D+1}}$.

It is clear that the life expectancy case is a special case of the life course perspective where $\bar{s} = 1$. From the result in Proposition 2 we see that the actuarially fair penalty schedule is strictly concave in age at benefit take-up in the more general case as well. As we show in Appendix B, the penalty schedule decreases as the survival probability, $S(a)$, increases.[22] As in the case of an increase in life expectancy above, the intuition is that a person who lives longer will require a lower penalty for early benefit take-up to be made indifferent because the same rate of (permanent) benefit reduction causes a greater lifetime loss. Also as before, the effect of a higher discount rate is to lower the benefit schedule and increase its curvature.

4.2 Assessment

The key result of the theoretical analysis above is that, under constant discounting, the penalty schedule that makes the average mortality individual indifferent between claiming before NRA and at NRA is decreasing at an increasing rate (penalty schedule is strictly concave) in take-up age before NRA. In contrast, the penalty schedule used by Social Security up to the 1937 cohort is linear in take-up age as shown in Figure 4. For later cohorts facing higher NRAs, the benefit schedule is piece-wise linear between with a smaller initial marginal penalty between ERA and NRA - 3 Years than after that, implying a more concave penalty schedule overall as discussed above. This suggests that none of the penalty schedules found in Social Security law satisfies the actuarial fairness principle completely under constant discounting but current and future cohorts with higher NRA face schedules that are more closely aligned with the actuarially fair form. In the remainder, our focus is on how far the

[22] It is easy to show that in the case of a constant survival rate, $S(a) = \bar{s}$, the penalty schedule at age 64 decreases if and only if $\frac{1-(\bar{s}\beta)^D}{(1-\bar{s}\beta)(\bar{s}\beta)^D} \geq D$. This condition holds for any $0 \leq \bar{s} \leq 1$ and $0 \leq \beta \leq 1$ if $D \geq 1$.

actual penalty schedules are from being actuarially fair across cohorts and under different discount rate scenarios. We also investigate the sensitivity to different assumptions regarding future mortality and counterfactual retirement age policies.

Life Expectancy Perspective

We can illustrate the deviation of the actual penalty schedule from its fair form adopting the life expectancy perspective (Definition 1) using calibrations as shown in Table 1. Here we consider the cases of an NRA of 65, life expectancy of 79 and 82 years (conditional on reaching age 62), and three different discount rates corresponding to, respectively, no discounting ($\delta = 0.00$), moderate discounting ($\delta = 0.03$), and high discounting ($\delta = 0.09$). The intermediate value of an annual (real) discount rate of 3% is chosen to be roughly consistent with the assumptions by the Actuarial Office of Social Security based on the real rate of return on (risk-free) assets (see e.g., Girola, 2005).[23] The high discount rate scenario is chosen to illustrate the perspective of individuals whose subjective time preferences significantly exceed the opportunity cost of time implied by the real interest rate.[24] The life expectancies are chosen to roughly match the difference in average old-age life expectancy between the 1960s and the 1990s (see Figure 2), the period between the introduction of the early retirement option and just prior to the first major adjustment of the early retirement penalty schedule (increase in NRA from 65 to 66 after 1999). Table 1 also displays the actual penalty schedule when benefits are initiated at exact ages 62, 63, 64, and 65 based on NRA 65 $(0.200, 0.133, 0.067, 0.000)$, that is the pre 1983 reform schedule that was in effect until 1999. The results in Table 1 confirm that the fair penalty schedule is decreasing at an increasing rate (strictly concave) under all discount rate scenarios. It also illustrates the effect of discounting on the benefit schedule: higher discounting results in a steeper penalty schedule, all else

[23]As noted by Girola (2005) and Sun and Webb (2009), 3% has exceeded the real risk-free interest rate implied by long-dated TIPS in recent years. Based on our own calculations we estimate an average annual real interest rate on 20 year treasuries of 2.76% based on 1954-2010 data (see Figure 5).

[24]For example, Gustman and Steinmeier (2002) estimate a structural life cycle model of retirement using data from the Health and Retirement Study. They suggest that empirical time preference rate follows a bimodal distribution such that they are either very low or very high. At the low end, nearly 40% of the sample, according to their estimates, have time preference rates below 5% and 21% have rates between 5% and 10%. At the upper tail of the distribution, 27% of the sample have discount rates of 50% or greater. It appears that less than half of the individuals in their sample have a discount rate in line with the risk-free real interest rate.

equal, as more impatient individuals require a greater penalty to be made indifferent between reduced early and normal benefits compared to more patient individuals. For 82 years of life expectancy, the penalty under no discounting is $(0.150, 0.105, 0.056, 0.000)$ compared to $(0.190, 0.134, 0.071, 0.000)$ under moderate discounting and $(0.277, 0.197, 0.105, 0.000)$ under high discounting. The results show that a smaller penalty rate is required for the higher life expectancy, because beneficiaries will receive reduced benefits for a longer period of time.

Comparing the fair schedules for the different discount rate scenarios to the actual schedule provides a sense of how far the actual schedules are from being actuarially fair. Under no discounting, the actual penalties are generally too high to be actuarially fair. In the high discounting case ($\delta = 0.09$), the penalty levels would have to be increased significantly from their current levels. For a moderate discount rate of 0.03, a value that is consistent with the long-term average real annual rate of return on risk-free assets, the constructed penalty schedule matches the actual schedule used by Social Security fairly closely. We can see that for this level of time discounting, the pre 1983 reform penalty schedule (NRA 65) comes closest to a fair schedule for a life expectancy of 82 years, that is for cohorts that reached their early retirement age in the 1990s, just prior to the gradual increase in the NRA from 65 to 66. Higher life expectancy also results in the fair penalty schedule to be more concave (more curved), facilitating greater resemblance with the actual schedule that retirees faced until 1999. The figures in Table 1 show that for the moderate discount rate of 3%, a level most consistent with values used by Social Security actuaries, the actual penalty schedule is far from the actuarially fair schedule for beneficiaries with expected length of life of around 79 years (matching those who reached their retirement age in the 1960s or 1970s). Taken at face value, this implies that the actual early take-up penalties are relatively too high for these individuals.

A related interest is the (constant) discount rate that makes a penalty schedule closest to actuarially fair for a given life expectancy. This question is of interest since life expectancy has been rising continuously (see Figure 2) while Social Security did not adjust the penalty schedule until 2000 when the 1983 Amendments started to take their effects on the NRA. We calculated the "best fit discount rate" which is the rate that minimizes the sum of the squared distances between the actual NRA 65 penalty rates and the fair penalty rates for a given life expectancy (conditional on reaching age 62) at ages 62, 63, and 64. Figure 6 shows the relationship between life expectancy and that discount rate. The positive relationship reflects the fact that at the same penalty schedule an individual whose life expectancy increases requires that future payments be more heavily discounted to remain indifferent between early and normal benefits. Specifically, the implied best fit discount rate for the 1908 birth cohort which had a life expectancy of 79.1 years when it reached 62 in 1970 is 1.85%. The 1918 birth cohort had a life expectancy of 80.6 years when it reached 62 in 1980 with an implied discount rate of 2.85%. The corresponding discount rate of the last cohort to face the NRA 65 penalty schedule, the 1937 cohort which had a life expectancy of 82.2 years when it reached 62 in 1999, is 3.7%.

Social Security actuaries claimed in the 1990s that rising real interest rates had effectively offset rising longevity such that the early benefits of the NRA 65 schedule continued to be roughly actuarially fair for the cohorts reaching retirement age at that time.[25,26] As shown in Figure 5, the long-term trend in the real interest rate 1954-2010 is positive but the series is quite volatile (a linear time trend only explains 6.7% of the variation in the interest rate) and the trend reversed to negative in the mid 1980s. Based on the trendline and the life expectancy we can calculate an average value of the real interest rate between age 62 and the mean age at death. For the 1908 birth cohort this interest rate is 2.65%.[27] It increases to 3.02% for the 1918 cohort, 3.38% for the 1928 cohort, and 3.7% for the 1937 cohort. This 0.73 percentage point increase between the 1908 cohort and the 1928 cohort compares to a 1.55 percentage point increase required to maintain actuarial fairness for an increase in life expectancy corresponding to that observed between these cohorts based on our implied best fit discount

[25]Professor Reimers told us about an exchange she had with Social Security in the 1990s: *Back in the 1990s I asked the actuaries at the Social Security Administration how they could continue to claim the benefit adjustment for early retirement was "actuarially fair," given the increase in longevity. Their answer was that the trend in interest rates had roughly offset the trend in longevity.*

[26]Using unpublished interest rate and life expectancy data from the Office of the Actuary of the Social Security Administration, a 2004 policy brief by Jivan (2004, p.2, Figure 3) reports that the ratio of the present values of the benefit stream for take-up at age 62 to take-up at 65 was about one, a value referred to as indicating "perfect age neutrality" of the 20% penalty for take-up at exact age 62, in the 1960s and has returned to levels close to that in the 1990s. The discount rate in the present value calculation is based on a constant real interest rate but it is not clear what the underlying average interest rates are. The two real interest rates reported are for the years 1960 and 2004, 1.4% and 3%, and do not match up with our values for these years, 2.34% and 2.38%, as shown in Figure 5.

[27]This number is the average of 2.35% and 2.94%, the values on the trendline at age 62 and age 79.1. The latter is the life expectancy conditional on survival to age 62 in 1970 from period life table data.

rate calculation discussed above (see Figure 6). (Looking at relative rather than absolute changes, the required increase in the implied discount rate is even more dramatic.) These back-of-the-envelope calculations suggest that the positive trend in real interest rates provided some actuarial offset to the gains in longevity across cohorts facing NRA 65. However, any compensation from higher interest rates was temporary since the trendline provided an increasingly poor fit after the mid 1980s when interest rates started a two-decade-long decline.

Life Course Perspective

We now turn to the life course perspective (Definition 2) which enables us to accurately incorporate differences in age-specific mortality past age 62 across cohorts. Using cohort-level data on average annual age-specific probabilities of dying, we calculate expected present values as of age 62 of the lifetime Social Security benefit stream associated with take-up at exact age 62, 63, ..., NRA. For each cohort, we compare the expected present values of claiming before the NRA to claiming at NRA. The difference between the expected benefits when claiming at say age 62 versus at NRA is a measure of how close the early retirement benefit schedule is to being actuarially fair for a given cohort. As discussed above, we obtained the survival probabilities from Cohort Life Table data.[28] In addition to the moderate discounting (constant 3%) and the high discounting (constant 9%) scenarios considered above, we also consider a cohort-specific variable discounting scenario, consistent with the evolution of real interest rates over the cohort's lifetime past age 62. Annual real benefits at NRA are normalized to \$1. (For comparison, the average PIA of retired workers age 66 in December 2009 was \$1,414.90 according to SSA-S 2010 Table 5.B2, suggesting that average annual benefits of Americans who turned 66 and claimed benefits that month are about \$16,979 in current dollars.)

Analysis of Present Value Graphs

Figures 7-10 plot the expected present values of the benefit stream by age at take-up for selected co-

[28]Age-specific cohort mortality rates after 2005 are not available. We impute missing data using the mortality rates of older cohort in 2005. This assumes that age-specific mortality remains at its 2005 level. To the extent that this imputation fails to account for future mortality declines at ages 62 to 110, the present values will be biased downwards. We examine the sensitivity of the results to this assumption further, using mortality rates predicted from models estimated from historic cohort data as discussed in more detail below.

horts, assuming moderate discounting (Figure 7), high discounting (Figure 8), current real discounting (Figure 9), and projected real discounting (Figure 10). The vertical scale is in dollar units where \$1 represents one year of full (normalized) benefits (benefits claimed at NRA). Our main interest is in the shape of the present value function, that is how the present values for take-up before NRA compare to the present value for take-up at NRA. It is also meaningful to compare present values at the same take-up age across cohorts. Recall that such differences can only be due to differences in the penalty schedule, the average mortality risk profile, and the discount rate. Cohorts are identified by birth year and, alternatively, by year in which individuals reach age 62 and become eligible for early benefits. A penalty schedule that is (perfectly) actuarially fair for a cohort would be depicted by a flat graph of the present value function. Differences in the present values imply that the underlying penalty schedule is not actuarially fair individuals with average age-specific mortality (for given discount rates). Present values for early claiming ages that are below those for take-up at NRA imply that early take-up is actuarially unfavorable relative to take-up at NRA. Similarly, when present values for early take-up exceed those at NRA then the penalties for early take-up are too low to be actuarially fair, providing average individuals with an incentive to take up benefits before reaching their NRA.

Looking at the four birth cohorts 1908, 1918, 1928 and 1937, all subject to NRA 65 and reaching age 62 in 1970, 1980, 1990 and 1999, respectively, we find that the earlier cohorts have lower present values as shown in Figures 7-10, reflecting their shorter average life spans. Comparing Figures 7 and 8, the graphs of the present value functions for 3% discount rate are fairly flat while the graphs for the 9% discount rate slope noticeably downward. This confirms that the NRA 65 schedule is closer to the actuarially fair penalty schedule in the case of moderate discounting. When future benefits are discounted more heavily, as in the 9% discount rate case shown, the penalties for early take-up are too low to make the average-mortality person in these cohorts indifferent between early (reduced) and normal (full) benefits. As shown in Figure 9, the present value graphs are similar to the moderate discounting case when we assume variable discounting based on period interest rates.[29] This is not

[29]The values for the periods after 2010 are imputed using the long-term average value of 2.76%. This closely matches the assumptions by Social Security in their long-term solvency calculations (intermediate case) where real interest rates stabilize at 2.9% in 2020 (SSA-T, 2009, p.101).

surprising since the average return over the period is 2.76%, close to our 3% baseline case.

For cohorts 1943 and 1960 that face NRA 66 and 67, respectively, the present value graphs are noticeably below those of the 1928. They also tend to be below the graphs of the 1918 cohort. This pattern is a reflection of the *ceteris paribus* reduction in lifetime benefit wealth associated with the higher NRA (see Figures 7 and 9). A gradual transition from NRA 65 to 66 (see Table 1) can be observed by following the graphs of the 1937-1943 cohorts.[30] Close inspection reveals that the present values for early claiming ages are more similar to the values at NRA for the 1918 and 1928 cohorts (NRA at exact age 65) than for the other cohorts. This suggests that Americans in these birth cohorts face a penalty schedule that is most closely aligned with the actuarially fair schedule implied by the life course perspective. The graphs also illustrate the impact of the introduction of the 5/12% per month reduction for cohorts born after 1937 (reaching 62 after 1999). The present values graphs extend to the respective higher NRA and kinks resulting from the different marginal penalty rates around ages 63 (NRA 66) and 64 (NRA 67) are visible in most discounting scenarios.

Under moderate and variable discounting (using period interest rates), the present values for cohorts subject to NRA 65 are typically lower at ERA and NRA than for the intermediate take-up ages 63 and 64, resulting in a hump-shaped graph of the present value function. For the 1908 cohort the present value of the expected benefit stream is highest for take-up at age 63 and lowest at 65. Benefit take-up before NRA was beneficial—in an actuarial sense—from the average mortality individual's perspective for the 1908 and 1918 cohorts. However, for average individuals in later cohorts also subject to NRA 65, take-up at exact age 62 became actuarially unfavorable. They face a present value graph that is lowest at 62 and highest at age 64, the year before NRA. A similar pattern is evident for the cohorts that face higher NRAs. As shown in Figures 7 and 9, under moderate and variable real discounting, the average individual born in 1943 or 1960 achieves the lowest present value at 62 and the highest at 65 and 66, respectively.

Sensitivity to Mortality Assumption

Figure 11 shows the present value graphs under moderate discounting for two different age-specific

[30] We note that for the 1938-1942 cohorts the value for take-up at NRA is not displayed since the NRA is at an interim month.

mortality scenarios. The solid lines correspond to the values above that impute the missing death probabilities after 2005 using the data from adjacent (older) cohorts. The graphs in dotted lines utilize probabilities of dying after 2005 predicted from a model that we estimated from age-specific cohort mortality data.[31] The predicted probabilities for selected cohorts and ages are shown in Figure 3. The corresponding present value graphs tend to lie above those based on the imputed values. This reflects the fact that the predicted mortality risks extrapolate the trends in gains in survival rates observed across the 1900-1943 cohorts while the imputed values assume that mortality remains at its 2005 levels (see Figure 3). Longer life spans translate into greater present values of pension wealth at all claiming ages. The analysis shows that pension wealth levels of current and future cohorts of retirees can benefit substantially from reductions in future mortality. For example, the present values of the 1960 cohort are as much as 8% higher in the scenario that assumes that the gains in life expectancy past age 62 will continue (see Figure 11). Gains in survival rates tend to increase the present values for take-up at or close to NRA more, resulting in present value graphs for predicted mortality that are not only shifted upward but also noticeably tilted counterclockwise for the 1943 and 1960 cohort relative to the case with imputed mortality. As expected, the 1960 cohort is most affected; under the optimistic mortality scenario the present value for take-up at NRA exceeds that of take-up at 63, resulting in a greater incentive to claim past age 63 than in the absence of future reductions in mortality. The implication of future gains in mortality is that the NRA 67 penalty schedule will conform less to the actuarial fair form.

Quantifying Proximity to Fair Values

Table 2 summarizes the actual and fair penalty schedules for the birth cohorts 1908, 1918, 1928, 1937, 1943, and 1960 by discount rate scenario. The actuarially fair penalties are the benefit reductions that make the present values for claiming before NRA exactly equal to the NRA level. Thus, if Social Secu-

[31] Using data on annual age-specific probabilities of dying at ages 62-110 for birth cohorts 1900-1943, we estimated the following regression model: ${}_1q_a(c) = e^{\alpha(1+\gamma\cdot(c-1832))\cdot(110-a)^{\beta}+\varepsilon}$, where ${}_1q_a(c)$ is the probability of dying between exact ages a and a+1 for the average individual in birth cohort c. The MLE estimates obtained are (standard errors in parentheses): $\alpha = -0.12186$ (0.00086), $\beta = 0.81089$ (0.00192), and $\gamma = 0.00546$ (0.00009). For example, the model predicts that the probability of dying between ages 66 and 67 for a person born in 1960 is ${}_1q_{66} = e^{-0.12186\cdot(1+0.00546)\cdot(1960-1832)\cdot(110-66)^{0.81089}} = 0.01163$.

rity (had) adopted these penalty schedules, the corresponding present value graphs in the figures above would be exactly flat at the level of the present value for take-up at NRA. The present value graphs are a useful tool to visually assess how Social Security's penalty schedules compare to actuarially fair ones in the life course setting. To more accurately quantify how close a penalty schedule is to being actuarially fair, we also calculate an average proximity measure. Using the difference between the actual and the fair penalty schedules, we measure the degree of deviation from the actuarial fairness principle embedded in Social Security's schedules given mortality and discounting. Specifically, for each cohort we calculate the sum of the squared differences between the fair and the actual penalty over the exact early claiming ages. The last column in Table 2, under the heading of "Average Distance," reports the square-root of this measure divided by $NRA - ERA$ to account for differences in the number of early take-up ages. Values closer to 0 imply greater adherence to the actuarial fairness principle.

The quantitative analysis confirms that under moderate discounting (3%) Social Security's penalty schedule has been closest to actuarially fair for the 1918 and the 1928 birth cohorts (see Table 2). The average distance values for these cohort are 0.0026 and 0.0032, respectively, compared to 0.0046 for the 1908 birth cohort and 0.0044 for the 1937 cohort. The actual penalties for early take-up are generally too low for the 1908 cohort and too high for the 1937 cohort compared to the actuarially fair reductions. In particular, the penalty for take-up at ages 62 and 63 increased relative to the 1918 birth cohort as subsequent cohorts enjoyed reductions in age-specific mortality (see Figure 3). Under variable discounting adherence to actuarial fairness is very similar across the 1908, 1918, and 1928 cohorts. In this case the average-mortality individuals born in 1928 face penalties closest to the actuarially fair ones. The corresponding average distance value is 0.0050 suggesting a worse fit than under moderate discounting (see Table 9). As under moderate discounting, the fit between the actual penalties and the fair schedule worsens for later cohorts subject to NRA 65. For example, the implied fair penalty for Americans born in 1937 who claim at exact age 62 is 18.71% in the variable discount case, significantly below the 20% penalty that this cohort actually faced. Assuming moderate discounting, cohorts with an NRA of 66 (or greater) face early benefits that are actually more in line with actuarial fairness than the 1937 and 1928 cohorts. For example, as illustrated in Table 2, the average distance for the 1943 cohort is 0.0029 at 3% discounting compared to 0.0044 for the 1937 cohort, a 34% improvement. The planned increase in the NRA to age 67 over the next two decades is expected to have no worsening effect on compliance to actuarial fairness (0.0029 for the 1960 cohort). However, as shown in Figure 11, additional gains in life expectancy would result in significantly lower compliance (0.0075 for predicted mortality).

Our analysis of the early cohorts facing the NRA 65 penalty schedule, namely cohorts 1908, 1918, and 1928, is primarily ex-post since the majority of the members of these cohorts have completed their lifespans at this time. Analyzing recently implemented or proposed (counterfactual) penalty schedules in real time, like our assessment for cohorts subject to NRA 66 and 67, will necessarily be a forward-looking exercise and thus be reliant upon assumptions regarding the future paths of mortality and real interest rates. This raises the possibility that a penalty schedule found to deviate from actuarial fairness in an ex-post analysis, was (approximately) actuarially fair ex-ante, that is when designed or implemented. We will briefly examine this possibility, focusing on the early cohorts facing the NRA 65 penalty schedule, for which we observe almost completed life spans. As shown in Figures 2 and 3, mortality rates have declined gradually across periods/cohorts. Real interest rates, on the other hand, have been very volatile, making them considerably more difficult to predict (see discussion above) for the actuaries of Social Security. We consider the case where Social Security discounts using real interest rate values predicted from linear trendmodels that are estimated from the 15 years of data leading up to the period in which the cohort turns 62. The underlying trendlines for the main cohorts of interest are shown in Figure 5. The corresponding present value graphs in Figure 10 and the proximity results in Table 2 for the "15-year trend discounting" scenario show that, when real interest rates are extrapolated in this ex-ante way, the NRA 65 penalty schedule does not fit the fair schedules as well as under the current real interest rate or the constant 3% discounting scenario.[32] In particular, for the 1908 birth cohort, the cohort closest to the period when the ERA and the early take-up penalties were instituted, the value of the measure of average distance between actual and fair schedule is 0.0079 for 15-year trend discounting, compared to 0.0046 and 0.0051 for 3% constant discounting and current

[32]Notice that the scale in Figure 10 is much wider than in Figures 7-9, which gives the false impression that the present value graphs are flatter. This is the result of the greater variability in the average discount rates across cohorts.

real interest rate discounting, respectively. While we cannot rule out that the Social Security actuaries were considering other plausible interest rate scenarios as well when designing and implementing the NRA 65 penalty schedule, this does cast doubt at the idea that the deviations from actuarial fairness that we document in our (primarily) ex-post analysis of the NRA 65 cohorts, are driven by real interest rate developments that were unanticipated at the time the NRA 65 penalty schedule was implemented. Another way of looking at this result is that the NRA 65 penalty schedule may have turned out to be closer to actuarially fair than originally anticipated or planned by Social Security. This is consistent with the fact that the NRA 65 schedule is closer to actuarially fair for the 1918 and 1928 cohorts than for the 1908 cohort under both, period interest discounting as well as 3% discounting (which is above the 1955-79 average as shown in Figure 5). As discussed above, the difficult-to-predict long-term positive trend in the real interest rate actuarially balanced, to some extent, the gradual, and hence more likely anticipated, reductions in mortality rates.

Retirement Age Reform and the Penalty Schedule

The discussion in Myers and Schobel (1990, Footnote 3, p.296) suggests that Congress, at the time of the deliberations for the 1983 Amendments, also considered a gradual increase of the ERA in tandem with the NRA. Had this been implemented, the ERA for the 1943 and 1960 cohorts would have been 63 and 64, respectively. In that case there would have been no need to introduce the 5/12 of 1% reduction per month (which was actuarially too small) and the slope of the penalty schedules faced by the 1943 and the 1960 cohorts would have been the same 5/9 of 1% reduction per month. Figure 12 shows the present value graphs for the 1943 and the 1960 cohorts under this counterfactual retirement age reform scenarios for both imputed and predicted mortality rates and assuming a discount rate of 3%. We note that this would have resulted in a more severe reduction in benefits than the reforms that were actually passed (e.g., 20% reduction at age 64 compared to 16.67% under the current plan for those born in 1960 and thereafter). Comparing the average proximity values reported alongside the graphs in Figure 12 to the corresponding values in Table 2 (1943 and 1960, 3%), it is clear that this gradual increase of the ERA in tandem with the NRA would have resulted in less deviation from actuarial fairness. For example, for the 1943 (1960) cohort and imputed mortality the proximity value for the counterfactual policy is 0.0032 (0.0032) compared to 0.0071 (0.0096) under current law. The conclusions are similar for imputed and predicted mortality as shown in Figure 12.

Increases in the NRA beyond age 67 are part of the reform scenarios considered in the current long-term solvency debate. It is clear that such reforms directly affect the penalty schedule and thus provide an opportunity to reassess compliance with the actuarial fairness principle. Figure 12 shows the results of two scenarios with NRA 69 and ERA 62 implemented for the 1960 cohort under 3% discounting. Given the popularity of early benefit take-up, increases in the NRA that maintain an ERA of age 62 are likely more politically feasible ways to achieve benefit reductions than reforms that increase the ERA. In the first case, we assume that Congress maintains the current marginal reductions of 5/9 of 1% per month for the period NRA - 3 Years and NRA, that is between 66 and 69 and 5/12 of 1% per month prior to that. The resulting penalties for take-up at ages 62, 65 and 67 are 40.0%, 26.67% and 13.33%. The proximity measure for this case under the conservative mortality scenario is 0.0038 and 0.0085 for our mortality projections. While the values are slightly higher than for NRA 67, it is clear that increases in the NRA in this fashion result in penalty schedules that are reasonable close to the actuarially fair form in the absence of further gains in mortality. In turn, this highlights the sensitivity of the results to the evolution of mortality, consistent with our findings above. We propose an alternative penalty formula that better adheres to the curvature requirement by penalizing early take-up more closer to NRA: 5/8 of 1% reduction per month between 68 and 69, 5/9 of 1% between 67 and 68, ..., and 5/14 of 1% between 62 and 63. For NRA 69 the resulting penalties for take-up at ages 62, 65 and 67 are 39.52%, 25.62% and 14.17% in this case. As shown in Figure 12, the corresponding present value graphs are much flatter than those that extrapolate the current penalty structure to NRA 69. The greater adherence to actuarial fairness of the 5/14-5/8 penalty schedule is also apparent in the significantly lower average distance values (imputed: 0.0006 vs. 0.0038; predicted: 0.0060 vs. 0.0085).

5 Conclusions

This paper provides a comprehensive analysis of the actuarial properties of the early retirement penalties of Social Security. The reductions for claiming benefits early are often taken to be actuarially fair, suggesting that, from the perspective of Social Security, the present value of the lifetime benefit stream for early take-up is the same as for take-up at NRA for average mortality Americans. We show formally that, for a constant discount rate, actuarial fairness requires the penalty schedule to be decreasing at an increasing rate in age at early take-up. In contrast, the schedule used by Social Security until 1999 was decreasing at a constant rate, suggesting that it was never designed to be exactly actuarially fair. A preference for simplicity of the penalty schedule by Congress may have led to the initial linear penalty schedule (NRA 65). Under the assumption of moderate levels of time discounting (a constant 3%), consistent with rates used by Social Security actuaries based on average real returns on risk-free assets, the NRA 65 penalty schedule is found to be closest to actuarially fair for Americans with average mortality who reached retirement age in the 1980s. We also show that for cohorts subject to NRA 65, the positive trend in the real interest rate until the mid-1980s actuarially balanced, to some extent, the gains in life expectancy. An expectation of continuing increases in the real interest rate (which did not materialize) may have prevented the introduction of a steeper penalty schedule between 65 and NRA in the 1983 reforms. We project that the increases in the NRA from age 65 to 66 in recent years and the planned rise in the NRA to age 67 by 2020, following the 1983 Amendments, will result in penalty schedules that are more in line again with the actuarially fair from for average Americans.

These findings have important implications for the debate over the long-term solvency of Social Security. If the penalty schedules are approximately actuarially fair under the assumptions of the Social Security actuaries, then the total (expected) outlays for the average beneficiary are independent of the timing of benefit take-up. We found benefit take-up at any age before NRA to be beneficial for the average mortality individuals when moderate discount rates were assumed, for the 1908 and 1918 cohorts but for later cohorts also subject to NRA 65, take-up at exact age 62 became actuarially less favorable. The implied optimal age to claim benefits for average Americans facing NRA 65 was shown to be between age 63 and 64. Current average mortality retirees facing NRA 66 should be approximately indifferent between take-up at an early age and take-up at NRA, assuming moderate discounting. Those expecting to live longer than the average person in their cohort should claim later. As shown in Figure 13, our estimates based on Social Security's Public-Use Master Beneficiary Microdata suggest that the proportion of beneficiaries who claimed retirement benefits in the 12-months period after turning 62 increased from about 43% to over 50% across the 1911 to 1937 birth cohorts.[33] The period data reported in Table 3 suggests that the popularity of claiming as early as possible has remained high even after the increase in the NRA from 65 to 66.[34] Given the actuarially optimal claiming ages, the Social Security Trust Fund likely benefitted from the rise in early benefit take-up. While there is a sizable share of Americans claiming around age 65 which is optimal for the average person, the median person claims at 62 which should not contribute to the long-term balance between expected inflows and outlays from Social Security's lifetime budgetary perspective according to our calculations. However, there are likely some losses in Social Security payroll tax contributions as individuals claim and exit from the labor force earlier (see Benítez-Silva and Heiland, 2008). Also, any shift towards earlier take-up of benefits will (temporarily) put pressure on the cash-flows of the Social Security Trust Fund.

The analysis also provides insights into the optimal timing of benefit take-up across cohorts. Under the assumptions of moderate discounting, the optimal age at benefit take-up implied by our analysis for average Americans (Figures 7, 9, and 11) has risen from age 63 (1908 cohort) to age 65 (since the 1940 cohort). The continued high incidence of claiming at 62 (Figure 13 and Table 3) is remarkable in light of the improvements in old-age mortality and the changes in the penalty structure discussed here. Recent contributions to the life-cycle retirement literature have considered heterogeneity in subjective time preferences and benefit uncertainty as explanations for the high concentration of claiming at age 62. Gustman and Steinmeier (2002) capture important cross-sectional features of the claiming age distribution in a model allowing for a majority of individuals with discount rates above 5%. Benítez-Silva et al. (2009) estimate that 60% of Americans close to retirement age expect a cut in benefits

[33]This trend is less pronounced among female beneficiaries who traditionally claim earlier than men (see Benítez-Silva and Yin, 2009). Women face additional incentives to claim retirement benefits early (see Munnell and Soto, 2005).

[34]In 1994, the share of Americans claiming after their 63rd birthday and before turning 65 was 20%. In 2005, 57% of Americans claimed before their 63rd birthday, 38% initiated benefits between their 63rd and before their 66th birthday, and 5% claimed at exact age 66 or thereafter.

within the next ten years and show how rising benefit uncertainty can explain the ongoing popularity of benefit take-up at age 62. In our framework a rising likelihood that benefits will be cut would be reflected in rising subjective discount rates across cohorts. As shown at a discount rate of 9%, an empirically plausible level in light of the distribution of the discount rate reported by Gustman and Steinmeier (2002), the benefit schedule strongly favors take-up at age 62, the earliest possible age as shown in Figure 8. While the exact distribution of the subjective discount rate within and across cohorts is not known, our analysis can help address questions regarding the degree of impatience required to counteract the forces of mortality and early take-up penalty.

Our findings suggest that the recent increases and the scheduled future increases in the NRA provide Americans with a greater incentive to claim benefits close to their NRA. Further reductions in age-specific mortality after age 62 are likely and would add to this incentive. This development suggests that we may have seen the peak in take-up at age 62. In turn, increases in the working-lifetimes of these cohorts appear likely. Using the penalty schedule to influence working lifetimes was part of the motivations for the 1983 reforms. The debate at the time of the deliberations of the National Commission leading to the 1983 Amendments shows that the goal of affecting individuals' working-lifetimes was at the forefront of the discussion about the increases in the NRA and the associated rise in the early retirement penalty (see Myers, 1993, p.316). This idea of strictly penalizing early retirement lived on, both in the popular press and in the policy arena, when those reforms were proven to be insufficient to achieve long-term solvency of the system and left working lifetimes largely unaffected (see e.g., Brown 1996, p.32; Rejda 1999, p.112; World Bank 1994, pp.323-324). As we have argued here the 1983 Amendments did institute penalties that are just steep enough from an actuarial perspective in the absence of further gains in mortality. Additional increases in average life spans are likely and actuarial balance would require lower penalties. The current policy debate provides an opportunity to revisit this issue and reform the age-related incentives in the context of the goals of long-term budgetary soundness.

Appendix A

This Appendix defines the fair penalty schedule for risk-averse individuals. The basic life course case discussed in Section 3.2 considers risk-neutral individuals. Actual individuals will prefer a certain amount x over a lottery with expected value x. The fair penalty schedule for risk-averse individuals is defined as follows:
$\sum_{a=ERA}^{NRA+D}\prod_{i=ERA}^{a}S(i)\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot u((1-r(ERA))\cdot PIA+Y)=S(ERA)\cdot u(Y)+\sum_{a=ERA+1}^{NRA+D}\prod_{i=ERA}^{a}S(i)\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot u((1-r(ERA+1))\cdot PIA+Y)=...=S(ERA)\cdot u(Y)+S(ERA)\cdot S(ERA+1)\cdot\frac{1}{1+\delta}\cdot u(y)...+S(ERA)\cdot S(ERA+1)...\cdot S(NRA-1)\cdot(\frac{1}{1+\delta})^{NRA-ERA-1}\cdot u(Y)+\sum_{a=NRA}^{NRA+D}\prod_{i=ERA}^{a}S(i)\cdot(\frac{1}{1+\delta})^{a-ERA}\cdot u((1-r(NRA))\cdot PIA+Y)$
where $u(.)$ is a utility function and Y is a measure of income from other sources such as a private pension. In the life course model with risk-averse individuals whose (concave) preferences are summarized by a standard utility function $u(.)$, there is no closed-form solution for the penalty schedule. Comparative static analysis can be conducted using the implicit function theorem. The main properties of the fair benefit schedule shown in Proposition 2 in Section 4.1 are robust to how we model risk preferences. In particular, in the case of risk-averse individuals, the penalty schedule also decreases as the period survival probability increases. Similarly, the schedule rises as the time discount rate increases.

Appendix B

This Appendix provides a proof of a result stated in Proposition 2. We derive the effect of an increase in the period-survival risk, $S(j)$ ($j\in ERA,...,NRA+D$), on $r(c)$ ($c\in ERA,...,NRA-1$), the fair penalty rate at (early) retirement age c. Following Proposition 2, the fair penalty rate of individuals who claim at age c can be written as $r(c)=\frac{A}{A+B}$, where $A\equiv\sum_{a=c}^{NRA}\delta^{a-ERA}\prod_{i=ERA}^{a}S(i)>0$ and $B\equiv\sum_{a=NRA+1}^{NRA+D}\delta^{a-ERA}\prod_{i=ERA}^{a}S(i)>0$. Define $R\equiv\sum_{a=c}^{j-1}\delta^{a-ERA}\prod_{i=ERA}^{a}S(i)>0$ if $j>c$ and $R\equiv 0$ if $j\leq c$.

$\frac{dr(c)}{dS(j)}=\frac{\frac{dA}{dS(j)}B-A\frac{dB}{dS(j)}}{(A+B)^2}$

Case 1 ($j>NRA$):
$=\frac{0\cdot B-A\frac{B}{S(j)}}{(A+B)^2}=\frac{-A\frac{B}{S(j)}}{(A+B)^2}<0$

Case 2 ($j\leq NRA$):
$=\frac{\frac{dA}{dS(j)}B-A\frac{B}{S(j)}}{(A+B)^2}$ (since $\frac{dB}{dS(j)}=\sum_{a=NRA+1}^{NRA+D}\delta^{a-ERA}\prod_{i=ERA,i\neq j}^{a}S(i)=\frac{1}{S(j)}\sum_{a=NRA+1}^{NRA+D}\delta^{a-ERA}\prod_{i=ERA}^{a}S(i)$)

$=\frac{\frac{(A-R)B}{S(j)}-\frac{AB}{S(j)}}{(A+B)^2}$ (since $\frac{dA}{dS(j)}=\sum_{a=\max j,c}^{NRA}\delta^{a-ERA}\prod_{i=ERA,i\neq j}^{a}S(i)=\frac{1}{S(j)}\sum_{a=\max j,c}^{NRA}\delta^{a-ERA}\prod_{i=ERA}^{a}S(i)$
$=\frac{1}{S(j)}(A-R)$ with $R=\sum_{a=c}^{j-1}\delta^{a-ERA}\prod_{i=ERA}^{a}S(i)$ if $j>c$ and $R=0$ if $j\leq c$)
$=-\frac{B}{S(j)}R\leq 0$ (holds with strict inequality if $j>c$).

References

Auerbach, A., and K. Hassett (2006). "Optimal long-run fiscal policy: Constraints, preferences and the resolution of Uncertainty," *Journal of Economic Dynamics and Control,* **5** 1–22.

Baker, M., and D. Benjamin (1999). "How do retirement tests affect the labour supply of older men?" *Journal of Public Economics,* **71** 27–51.

Belloni, M., and C. Maccheroni (2006). "Actuarial Neutrality when Longevity Increases: An Application to the Italian Pension System," Center for Research on Pensions and Welfare Policies Working Paper 47/06.

Benítez-Silva, H., and F. Heiland (2007). "The Social Security Earnings Test and Work Incentives," *Journal of Policy Analysis and Management,* **26-3** 527–555.

Benítez-Silva, H., and F. Heiland (2008). "Early Claiming of Social Security Benefits and Labor Supply Behavior of Older Americans," *Applied Economics,* **40-23** 2969-2985.

Benítez-Silva, H., Dwyer, D., Heiland, F., and W. Sanderson (2009). "Retirement and Social Security Reform Expectations: A Solution to the New Early Retirement Puzzle," manuscript, SUNY-Stony Brook.

Benítez-Silva, H., and N. Yin (2009). "An Emirical Study of The Effects of Social Security Reforms on Benefit Claiming Behavior and Receipt Using Public-Use Administrative Microdata," *Social Security Bulletin,* **69-3** 77–95.

Blinder, A., Gordon, R., and D. Wise (1980). "Reconsidering the Work Disincentive Effects of Social Security," *National Tax Journal,* **33-4** 431–442.

Börsch-Supan, A. (2000). "A Model under Siege: A Case Study of the German Retirement Insurance System," *Economic Journal,* **110** F24–45.

Breyer, F., and S. Hupfeld (2007). "On The Fairness of Early Retirement Provisions," CESifoWorking Paper 2078.

Brown, R.L. (1996). "Social Security and Retirees: Two Views of the Projections, An Actuary's Perspective," in *Social Security: What Role for the Future?* P.A. Diamond, D.C. Lindeman, and H. Young (eds.) National Academy of Social Insurance. Washington, D.C.

Burkhauser, R.V. (1980). "The Early Acceptance of Social Security: An Asset Maximization Approach," *Industrial and Labor Relations Review,* **33-4** 484–492.

Burtless, G., and R. A. Moffitt (1985). "The Joint Choice of Retirement Age and Post-Retirement Hours of Work," *Journal of Labor Economics,* **3** 209–236.

Bütler, M. (1999). "Anticipation effects of looming public pension-reforms." *Carnegie-Rochester Conference Series on Public Policy,* **50** 119–159.

Coile, C., P. Diamond, J. Gruber, and A. Jousten (2002). "Delays in Claiming Social Security Benefits," *Journal of Public Economics,* **84-3** 357–385.

Crawford, V. P., and D. M. Lilien (1981). "Social Security and the Retirement Decision," *Quarterly Journal of Economics,* **96-3** 505–529.

Disney, R. and S. Smith (2001). “The labour supply effect of the abolition of the earnings rule for older workers in the United Kingdom,” *Economic Journal,* **112** C136–C152.

Duggan, J.E., and C.J. Soares (2002). “Actuarial Nonequivalence in Early and Delayed Social Security Benefit Claims,” *Public Finance Review,* **30(3)** 188–207.

Friedberg, L. (1998). “The Social Security Earnings Test and Labor Supply of Older Men,” in *Tax Policy and The Economy,* Vol. **12.** MIT Press.

Friedberg, L. (2000). “The Labor Supply Effects of the Social Security Earnings Test,” *Review of Economics and Statistics,* **82** 48–63.

Girola, J. (2005). “The Long-term Real Interest Rate for Social Security,” Department of Treasury Research Paper No. 2005-02.

Gruber, J., and P. Orszag (2003). “Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt?” *National Tax Journal,* **56(4)** 755–773.

Gruber, J., and D. Wise (2005). “Social Security Programs and Retirement around the World: Fiscal Implications,” NBER Working Paper No. 11290.

Gustman, A. L., and T. L. Steinmeier (1985). “Social Security Reform and Labor Supply,” NBER Working Paper No. 1212.

Gustman, A. L., & Steinmeier, T.L. (1991). “Changing the Social Security Rules for Work after 65,” *Industrial and Labor Relations Review,* **44** 733–745.

Gustman, A. L., and T. L. Steinmeier (2002). “The Social Security Early Entitlement Age in a Structural Model of Retirement and Wealth,” NBER Working Paper No. 9183.

Gustman, A. L., and T. L. Steinmeier (2004). “The Social Security Retirement Earnings Test, Retirement and Benefit Claiming,” Michigan Retirement Research Center No. 2004-090.

Honig, M., and C. Reimers (1989). “Is It Worth Eliminating the Retirement Test?” *American Economic Review,* **79** 103–107.

Hurd, M. (1990). “Research on the Elderly: Economic Status, Retirement, and Consumption and Saving,” *Journal of Economic Literature,* **28** 565–637.

Hurd, M., J. Smith, and J. Zissimopoulos (2002). “The Effects of Subjective Survival on Retirement and Social Security Claiming,” Michigan Retirement Research Center Working Paper 2002-021.

Jivan, N.A. (2004). “How Can the Actuarial Reduction for Social Security Early Retirement be Right?” *Just the Facts on Retirement Issues,* July 2004, Number 11, Center for Retirement Research at Boston College.

Leonesio, M.V. (1990). “Effects of the Social Security Earnings Test on the Labor Market Activity of Older Americans: A Review of the Evidence,” *Social Security Bulletin,* **53** 2–21.

Liu, L., and A. Rettenmaier (2006). “Work and Retirement,” The National Center for Policy Analysis Policy Backgrounder No. 162.

Lumsdaine, R. (1995). “Factors Affecting Labor Supply Decisions and Retirement Income,” NBER Working Paper No. 5223.

Lumsdaine, R., and O. Mitchell (1999). “New Developments in the Economic Analysis of Retirement,” in Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, Vol. **3C.** 3261–3307.

Mirer, T.W. (1998). “The Optimal Time to File for Social Security Benefits,” *Public Finance Review,* **26(6)** 611–636.

Munnell, A.H., and M. Soto (2005). “Why do Women Claim Social Security Benefits so Early?,” Center for Retirement Research at Boston College Working Paper 2005-35.

Munnell, A.H., S.A. Sass, A. Golub-Sass, and N. Karamcheva (2009). “Unusual Social Security Claiming Strategies: Costs and Distributional Effects,” Center for Retirement Research at Boston College Working Paper 2009-17.

Myers, R.J. (1993). *Social Security.* Fourth Edition. Pension Research Council and University of Pennsylvania Press.

Myers, R.J., and B.D. Schobel (1990). “Early-Retirement Reduction and Delayed-Retirement Increase Factors under U.S. Social Security Law,” *Transactions,* **XLII** 295–320.

Queisser, M, and E. Whitehouse (2006). “Neutral or Fair? Actuarial Concepts and Pension-System Design.” OECD Social Employment, and Migration Working Paper No. 40.

Reimers, C., and M. Honig (1993). “The Perceived Budget Constraint under Social Security: Evidence from Reentry Behavior,” *Journal of Labor Economics,* **11** 184–204.

Reimers, C., and M. Honig (1996). “Responses to Social Security by Men and Women: Myopic and Far-Sighted Behavior,” *Journal of Human Resources,* **31-2** 359–382.

Rejda, G. E. (1999). *Social Insurance and Economic Security.* Sixth Edition. Prentice Hall, New Jersey.

Ruhm, C. J. (1996). “Historical Trends and the Future of Older Americans,” in Crown, W.H. (Ed.), *The Handbook on Employment and the Elderly*, Greenwood.

Rust, J., and C. Phelan (1997). “How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets,” *Econometrica,* **65** 781–831.

Sun, W., and A. Webb (2009). “How Much do Households Really Lose by Claiming at 62?” Center for Retirement Research at Boston College Working Paper 2009-11.

SSA-S (various years). Annual Statistical Supplement to the Social Security Bulletin. `http://www.ssa.gov/policy/docs/statcomps/supplement/`

SSA-T (2009). *The 2009 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds.* Washington, D.C.

Votruba, M. E. (2003). “Social Security And Retirees’ Decision to Work,” manuscript, Case Western Reserve University.

Waldron, H. (2004). “Heterogeneity in Health and Mortality Risk Among Early Retiree Men,” Social Security Administration ORES Working Paper 105.

World Bank (1994). *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth.* Oxford University Press for the World Bank.

Table 1: **Actual vs. Fair Penalty Schedule (Average Life Expectancy Perspective, Definition 1)**

	Age at Benefit Take-up (NRA=65)			
	62	63	64	65
Actual Penalty	0.200	0.133	0.067	0.000
	No Discounting ($\delta = 0.00$)			
Mean Age at Death above Age 62 (M62) is 79 yrs	0.176	0.125	0.067	0.000
M62 is 82 yrs	0.150	0.105	0.056	0.000
	Moderate Discounting ($\delta = 0.03$)			
M62 is 79 yrs	0.215	0.152	0.081	0.000
M62 is 82 yrs	0.190	0.134	0.071	0.000
	High Discounting ($\delta = 0.09$)			
M62 is 79 yrs	0.296	0.212	0.114	0.000
M62 is 82 yrs	0.277	0.197	0.105	0.000

Notes: The penalty is the reduction in benefits (as a fraction of benefits at 65, NRA) by age of early take-up. The penalty is for take-up in the month the person turns 62, 63, 64 or 65. The mean age at death above age 62 (M62) is the mean length of life conditional on surviving to age 62.

Table 2: **Actual vs. Fair Penalty Schedule[a], Selected Cohorts**

	Age at Benefit Take-up					Average Distance
	62	63	64	65	66	$= \frac{\sqrt{\Sigma_{c=ERA}^{NRA-1}[r_{actual}(c)-r_{fair}(c)]^2}}{NRA-ERA}$
	Actual Penalty					
NRA=65	0.2000	0.1333	0.0667			
NRA=66	0.2500	0.2000	0.1333	0.0667		
NRA=67	0.3000	0.2500	0.2000	0.1333	0.0667	
	Actuarially Fair Penalty					**Distance**
1908 Birth Cohort (62 in 1970)						
3% Discounting ($\delta = 0.03$)	0.2046	0.1432	0.0753			0.0046
9% Discounting ($\delta = 0.09$)	0.3067	0.2181	0.1166			0.0484
Real Interest Period Discounting[b]	0.2057	0.1444	0.0757			0.0051
Real Interest 15-Year Trend Discounting[c]	0.1783	0.1242	0.0650			0.0079
1918 Birth Cohort (62 in 1980)						
3% Discounting ($\delta = 0.03$)	0.1972	0.1379	0.0725			0.0026
9% Discounting ($\delta = 0.09$)	0.3001	0.2131	0.1138			0.0455
Real Interest Period Discounting[b]	0.2084	0.1447	0.0739			0.0053
Real Interest 15-Year Trend Discounting[c]	0.0673	0.0458	0.0234			0.0549
1928 Birth Cohort (62 in 1990)						
3% Discounting ($\delta = 0.03$)	0.1910	0.1334	0.0700			0.0032
9% Discounting ($\delta = 0.09$)	0.2941	0.2086	0.1112			0.0428
Real Interest Period Discounting[b]	0.1859	0.1287	0.0666			0.0050
Real Interest 15-Year Trend Discounting[c]	0.3659	0.2680	0.1480			0.0762
1937 Birth Cohort (62 in 1999)[d]						
3% Discounting ($\delta = 0.03$)	0.1871	0.1306	0.0685			0.0044
9% Discounting ($\delta = 0.09$)	0.2898	0.2053	0.1093			0.0409
Real Interest Period Discounting[b]	0.1805	0.1258	0.0659			0.0070
Real Interest 15-Year Trend Discounting[c]	0.1098	0.0746	0.0380			0.0371
1943 Birth Cohort (62 in 2005)[e]						
3% Discounting ($\delta = 0.03$)	0.2436	0.1912	0.1337	0.0702		0.0029
9% Discounting ($\delta = 0.09$)	0.3677	0.2928	0.2077	0.1108		0.0433
Real Interest Period Discounting[b]	0.2432	0.1923	0.1358	0.0733		0.0031
Real Interest 15-Year Trend Discounting[c]	0.2025	0.1573	0.1088	0.0565		0.0173
1960 Birth Cohort (62 in 2022)[f]						
3% Discounting ($\delta = 0.03$)	0.2913	0.2436	0.1912	0.1337	0.0702	0.0029
9% Discounting ($\delta = 0.09$)	0.4337	0.3677	0.2928	0.2077	0.1108	0.0437
Real Interest Period Discounting[b]	0.2854	0.2385	0.1872	0.1308	0.0686	0.0046
Real Interest 15-Year Trend Discounting[c]	0.1751	0.1436	0.1105	0.0757	0.0389	0.0395

Notes: [a]The penalty schedule describes the relation between take-up age and benefit reduction. The *actuarially fair* penalty is the benefit reduction that forces the expected present value of the lifetime benefit stream from claiming early to be the same as at NRA (discounted to age 62) . [b]For the 1970-2010 period we use the actual annual real interest rates as reported in Figure 5. After 2010 we use a constant annual rate of 2.76% (the average over the 1954-2010 period). [c]For each cohort we extrapolate the annual real interest rates for ages 62-110 using the cohort-specific linear trends estimated from the 15 years prior to age 62 as shown in Figure 5. [d]Americans born in 1937 are the last cohort that faces an NRA of exactly 65 years. [e]Those born in 1943 are in the first cohort that faces an NRA of exactly 66 years. [f]The 1960 birth cohort is the first cohort facing an NRA of 67 years.

Table 3: **Social Security Claiming Behavior, 1994-2005. Proportions by age of first receipt.**

Age/Year	**1994**	**1995**	**1996**	**1997**[a]	**1998**[a]	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Age 62	0.5886	0.5825	0.6008	0.5968	0.5833	0.5858	0.5171	0.5539	0.5602	0.5699	0.5753	0.5663
Age 63	0.0789	0.0787	0.0746	0.0735	0.0801	0.0798	0.0671	0.0779	0.0777	0.0782	0.0810	0.0830
Age 64	0.1212	0.1160	0.1080	0.1046	0.1077	0.1077	0.1045	0.1344	0.1484	0.1273	0.1094	0.0992
Age 65	0.1566	0.1629	0.1568	0.1551	0.1557	0.1557	0.1959	0.1785	0.1724	0.1784	0.1862	0.1974
Age 66	0.0182	0.0178	0.0199	0.0210	0.0210	0.0194	0.0392	0.0130	0.0096	0.0105	0.0122	0.0146
Age 67-69	0.023	0.0245	0.0256	0.0339	0.0286	0.0291	0.0550	0.0199	0.0152	0.0160	0.0177	0.0187
Age 70+	0.0128	0.0171	0.0140	0.0147	0.0232	0.0221	0.0208	0.0221	0.0161	0.0193	0.0178	0.0204
# of Claimants[b]	*1,444.5*	*1,424.8*	*1,396.1*	*1,418.9*	*1,441.3*	*1,484.6*	*1,758.9*	*1,574.0*	*1,595.5*	*1,593.3*	*1,680.3*	*1,793.5*

Notes: [a]The percentages do not coincide with those reported in the Statistical Supplements since we have not counted the 120,000 widows who were converted in these years from widow benefits to retirement benefits. [b]In thousands of claimers. Does not include disability conversions at the NRA.

COHORT	NRA	YEAR AGE 62	PENALTY at 62	COHORT	NRA	YEAR AGE 62	PENALTY at 62	COHORT	NRA	YEAR AGE 62	PENALTY at 62
1908	65	1970	0.2000	1929	65	1991	0.2000	1950	66	2012	0.2500
1909	65	1971	0.2000	1930	65	1992	0.2000	1951	66	2013	0.2500
1910	65	1972	0.2000	1931	65	1993	0.2000	1952	66	2014	0.2500
1911	65	1973	0.2000	1932	65	1994	0.2000	1953	66	2015	0.2500
1912	65	1974	0.2000	1933	65	1995	0.2000	1954	66	2016	0.2500
1913	65	1975	0.2000	1934	65	1996	0.2000	1955	66+2mo	2017	0.2583
1914	65	1976	0.2000	1935	65	1997	0.2000	1956	66+4mo	2018	0.2667
1915	65	1977	0.2000	1936	65	1998	0.2000	1957	66+6mo	2019	0.2750
1916	65	1978	0.2000	1937	65	1999	0.2000	1958	66+8mo	2020	0.2833
1917	65	1979	0.2000	1938	65+2mo	2000	0.2083	1959	66+10mo	2021	0.2917
1918	65	1980	0.2000	1939	65+4mo	2001	0.2167	1960	67	2022	0.3000
1919	65	1981	0.2000	1940	65+6mo	2002	0.2250	1961	67	2023	0.3000
1920	65	1982	0.2000	1941	65+8mo	2003	0.2333	1962	67	2024	0.3000
1921	65	1983	0.2000	1942	65+10mo	2004	0.2417	1963	67	2025	0.3000
1922	65	1984	0.2000	1943	66	2005	0.2500	1964	67	2026	0.3000
1923	65	1985	0.2000	1944	66	2006	0.2500	1965	67	2027	0.3000
1924	65	1986	0.2000	1945	66	2007	0.2500	1966	67	2028	0.3000
1925	65	1987	0.2000	1946	66	2008	0.2500	1967	67	2029	0.3000
1926	65	1988	0.2000	1947	66	2009	0.2500	1968	67	2030	0.3000
1927	65	1989	0.2000	1948	66	2010	0.2500	1969	67	2031	0.3000
1928	65	1990	0.2000	1949	66	2011	0.2500	1970	67	2032	0.3000

Figure 1: Normal Retirement Age (NRA), Year Age 62, and Penalty Rate at 62 for Birth Cohorts 1908-1970

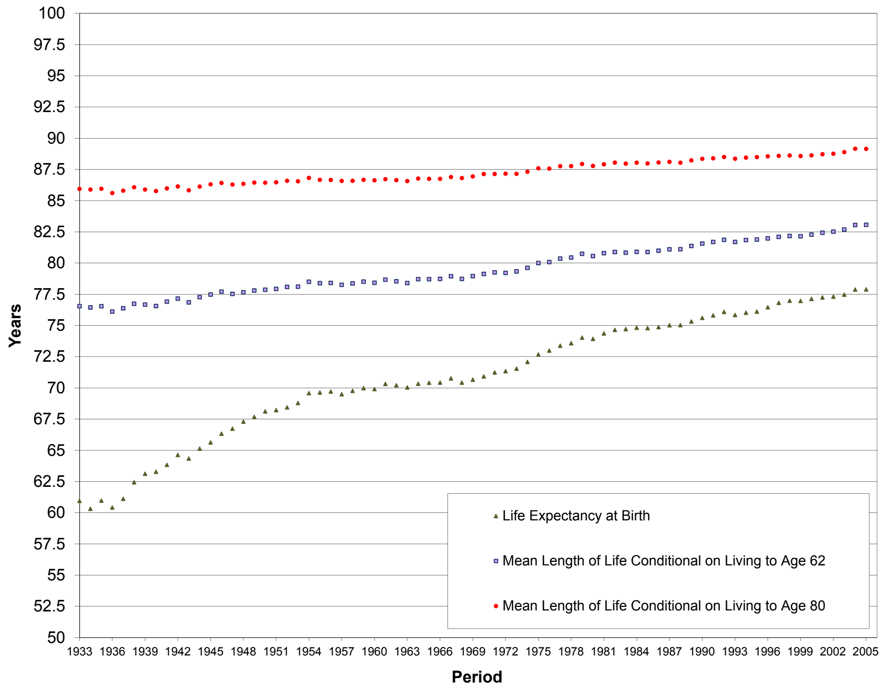

Figure 2: Life Expectancy at Birth, Mean Age at Death above Age 62 and above Age 80 (Source: Authors' calculations based on period life table data from the Human Mortality Database)

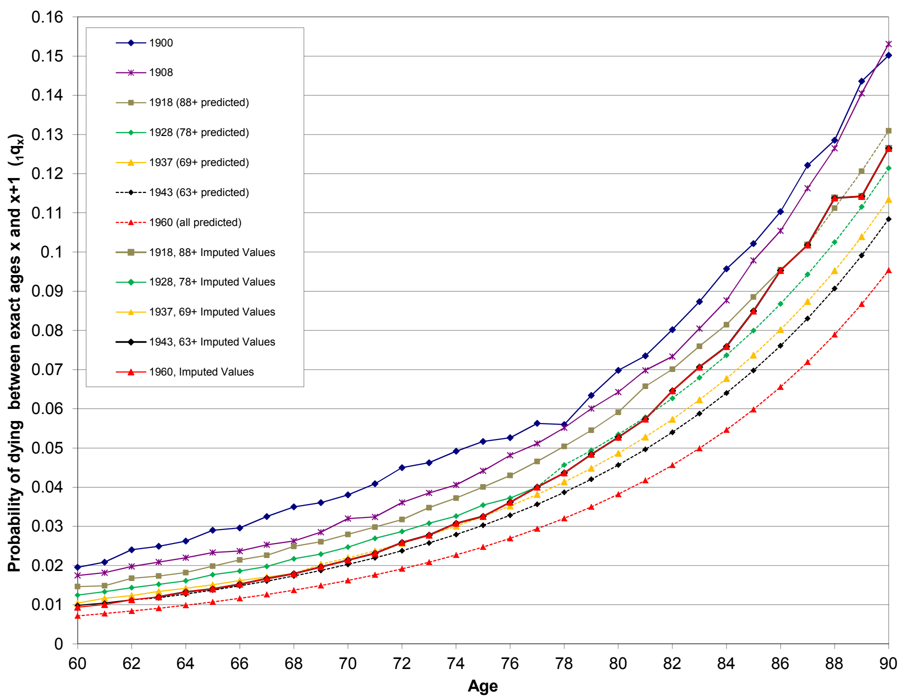

Figure 3: Age-Specific Probabilities of Dying, Age 60-90, Selected Cohorts 1900-1960 (Source: Authors' calculations and projections based on cohort life table data from the Human Mortality Database)

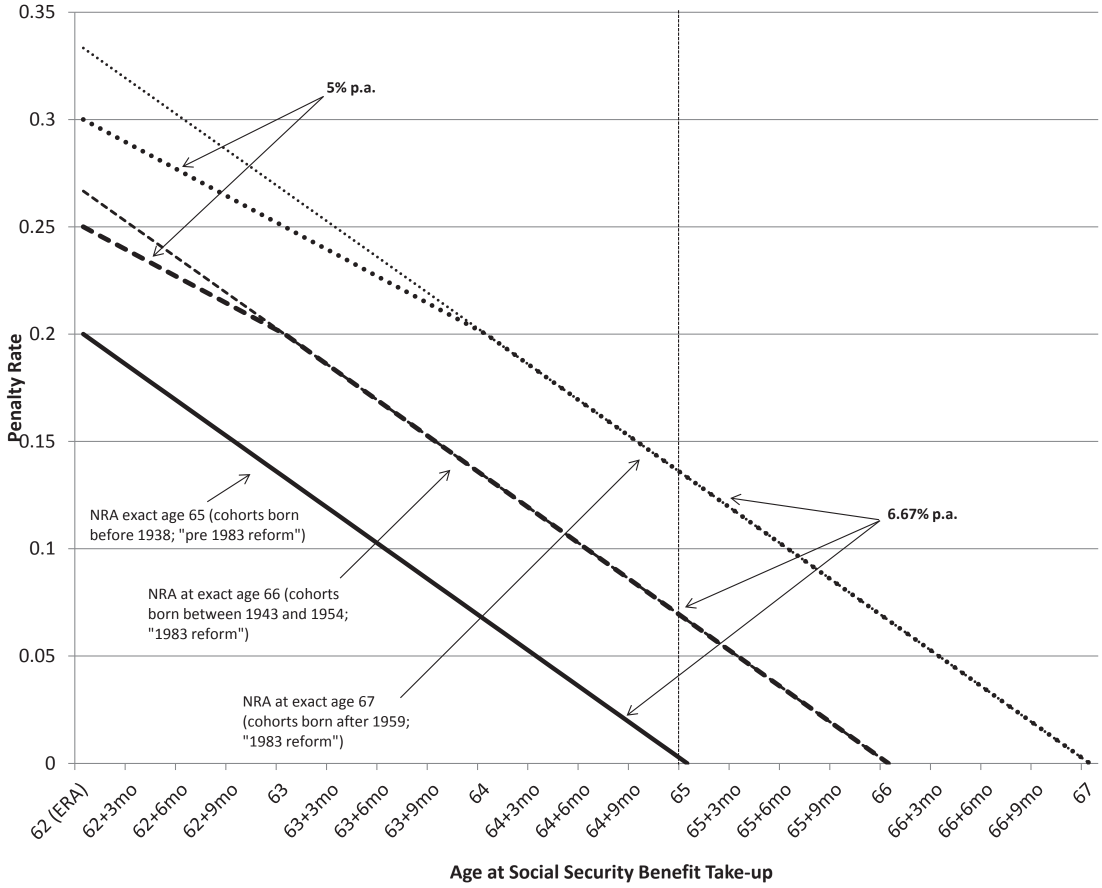

Figure 4: Early Retirement Penalty Rate Schedule

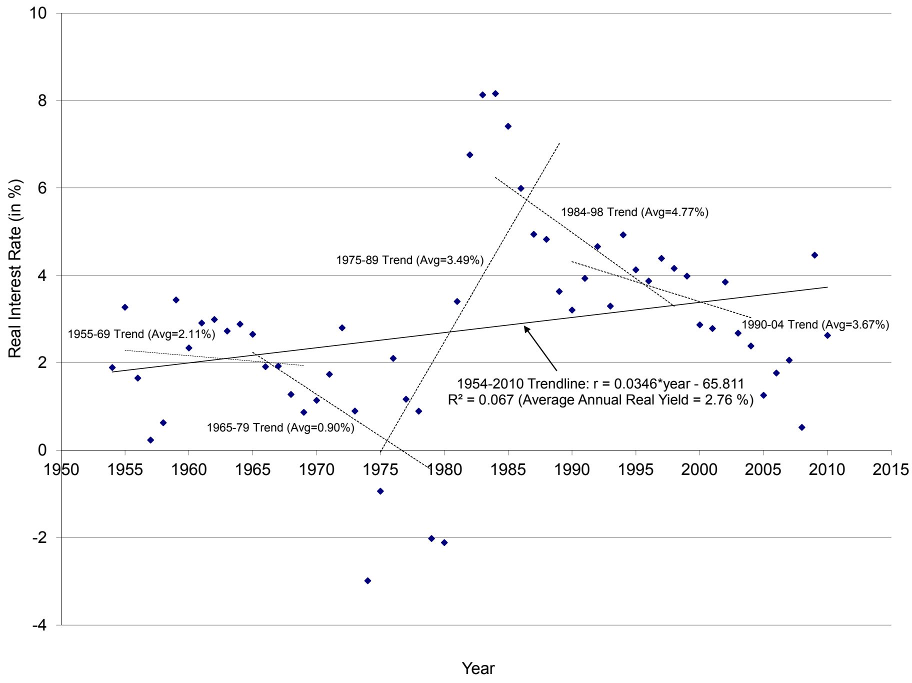

Figure 5: Annual Real Interest Rate (average annual yields of 20 year constant coupon treasuries net of inflation rate), Real Interest Rate 15-Year Linear Trends

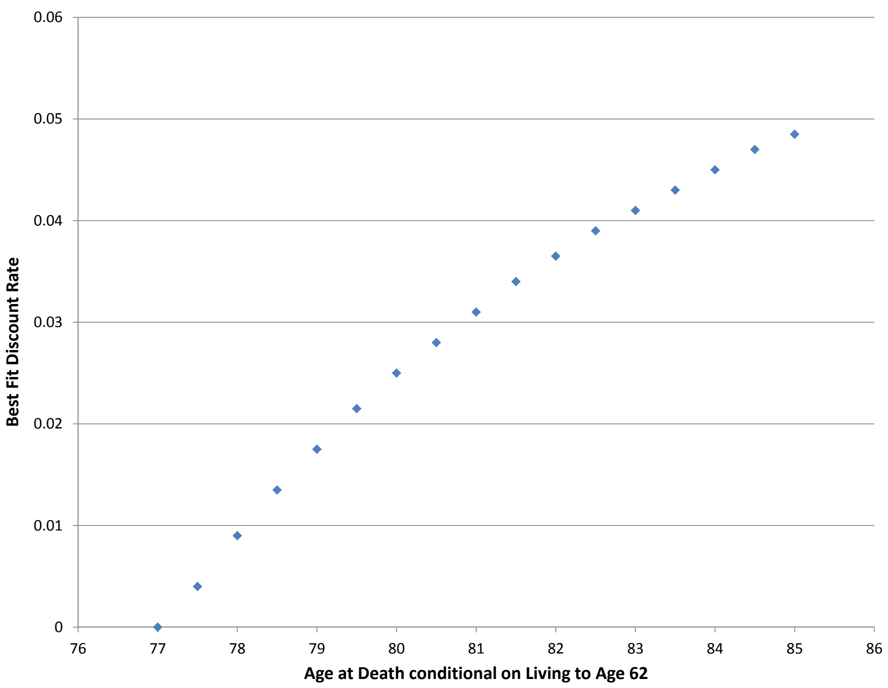

Figure 6: Combinations of Discount Rate and Life Expectancy such that NRA 65 Penalty Schedule is closest to Actuarially Fair (Life Expectancy Perspective)

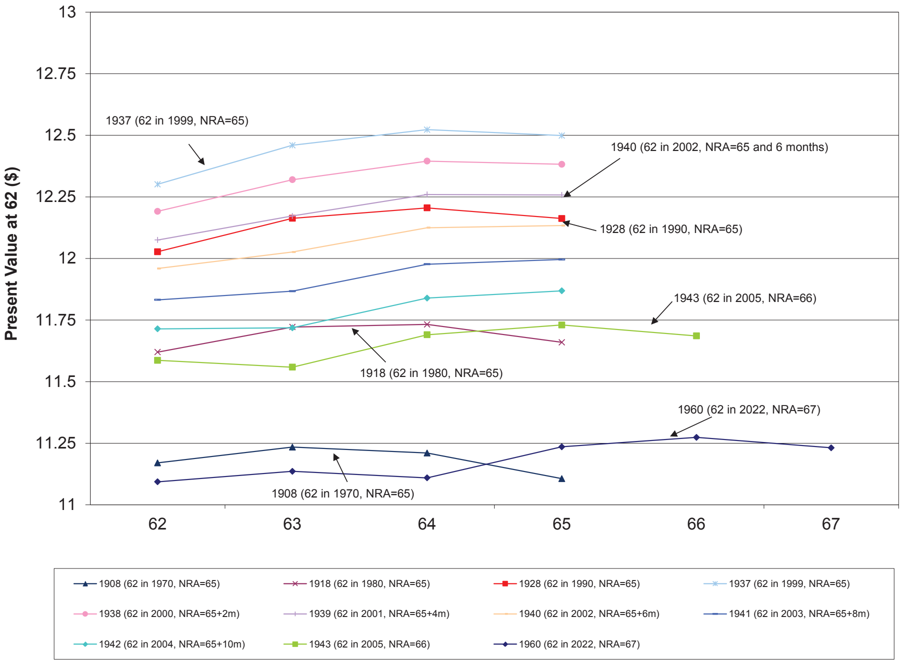

Figure 7: Present Value of Benefit Stream by Take-up Age, 3% Discount Rate

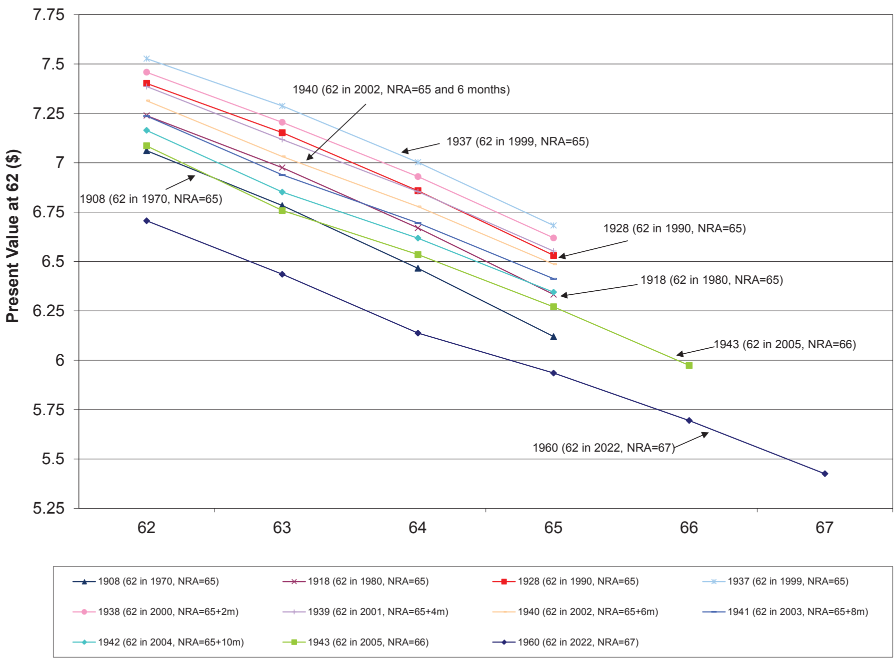

Figure 8: Present Value of Benefit Stream by Take-up Age, 9% Discount Rate

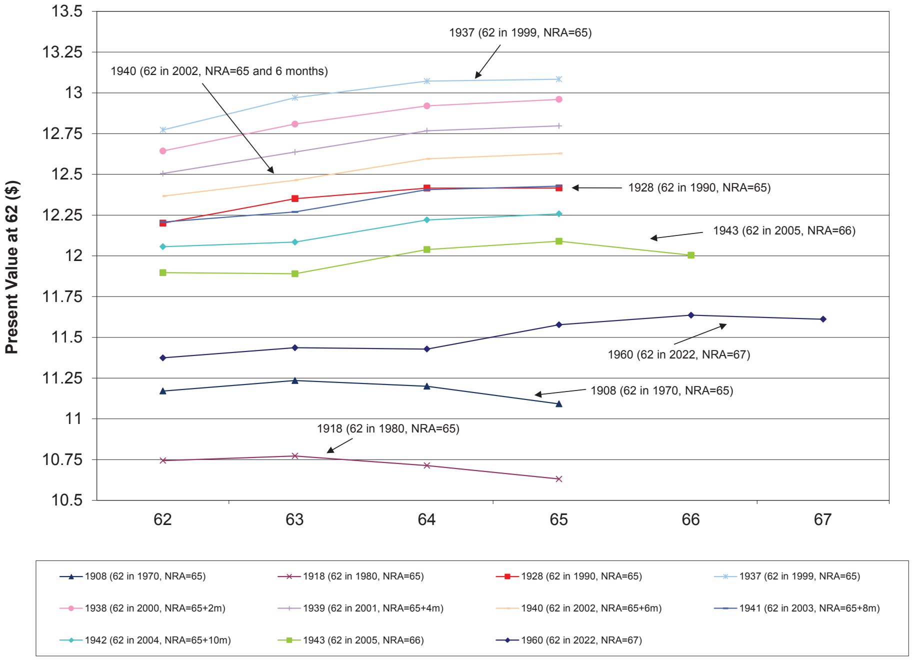

Figure 9: Present Value of Benefit Stream by Take-up Age, Current Annual Real Interest Rate Discounting

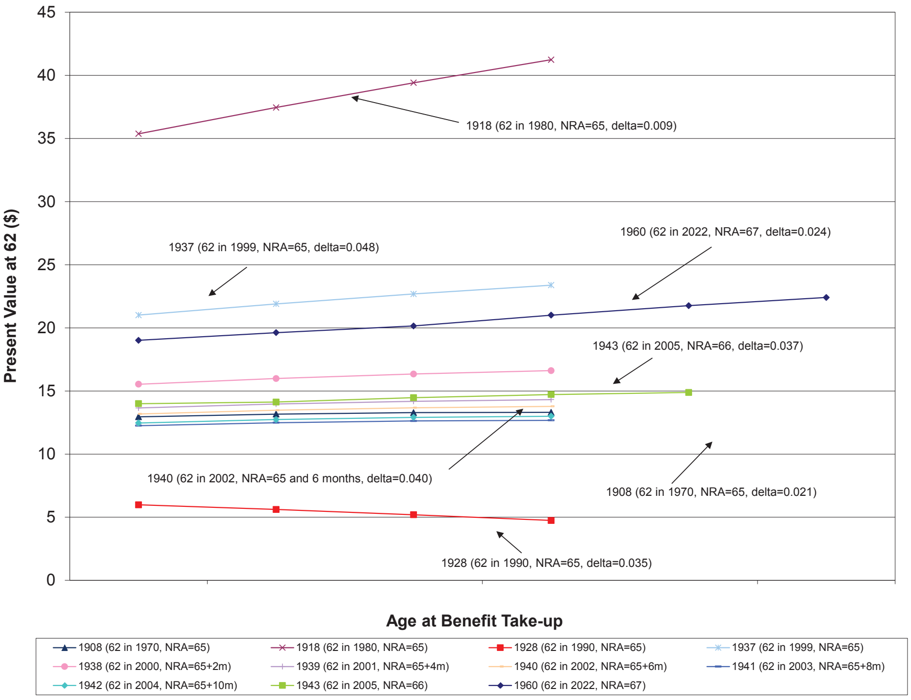

Figure 10: Present Value of Benefit Stream by Take-up Age, Real Interest Rate 15-Year Linear Trend Model Discounting

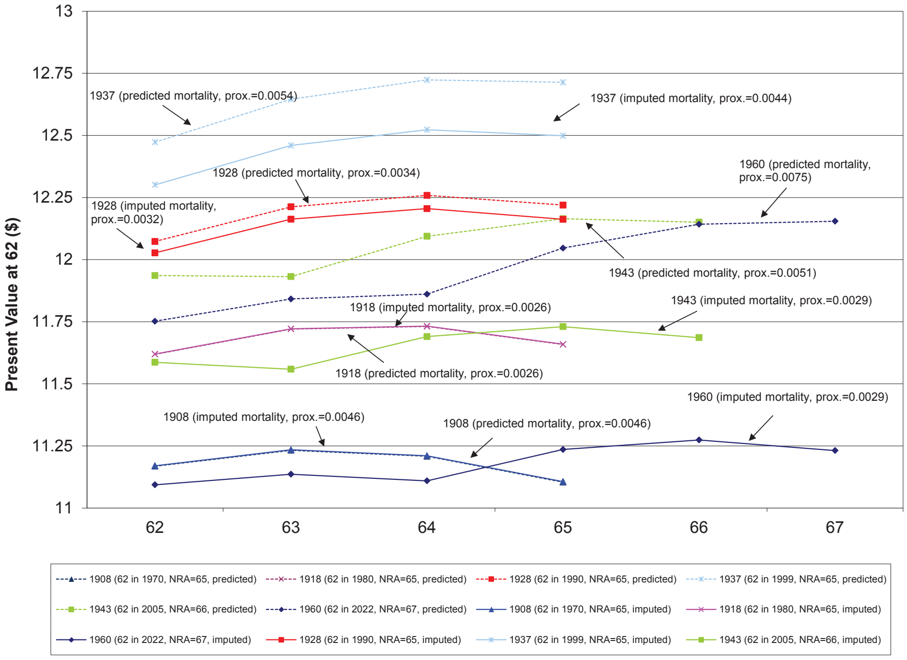

Figure 11: Sensitivity to Mortality Assumptions, 3% Discount Rate

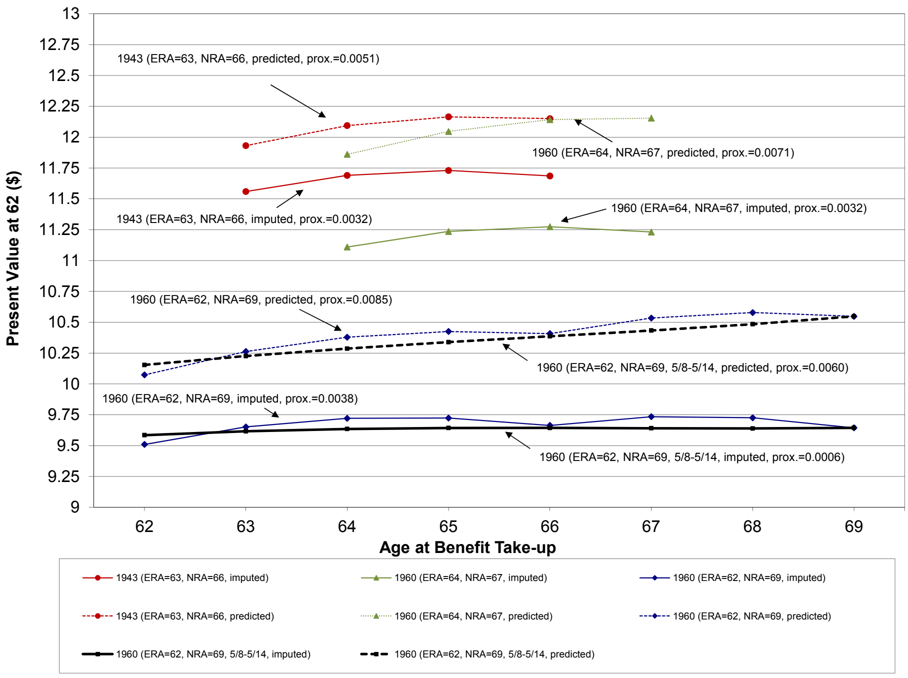

Figure 12: Retirement Age Reform Scenarios, 3% Discount Rate

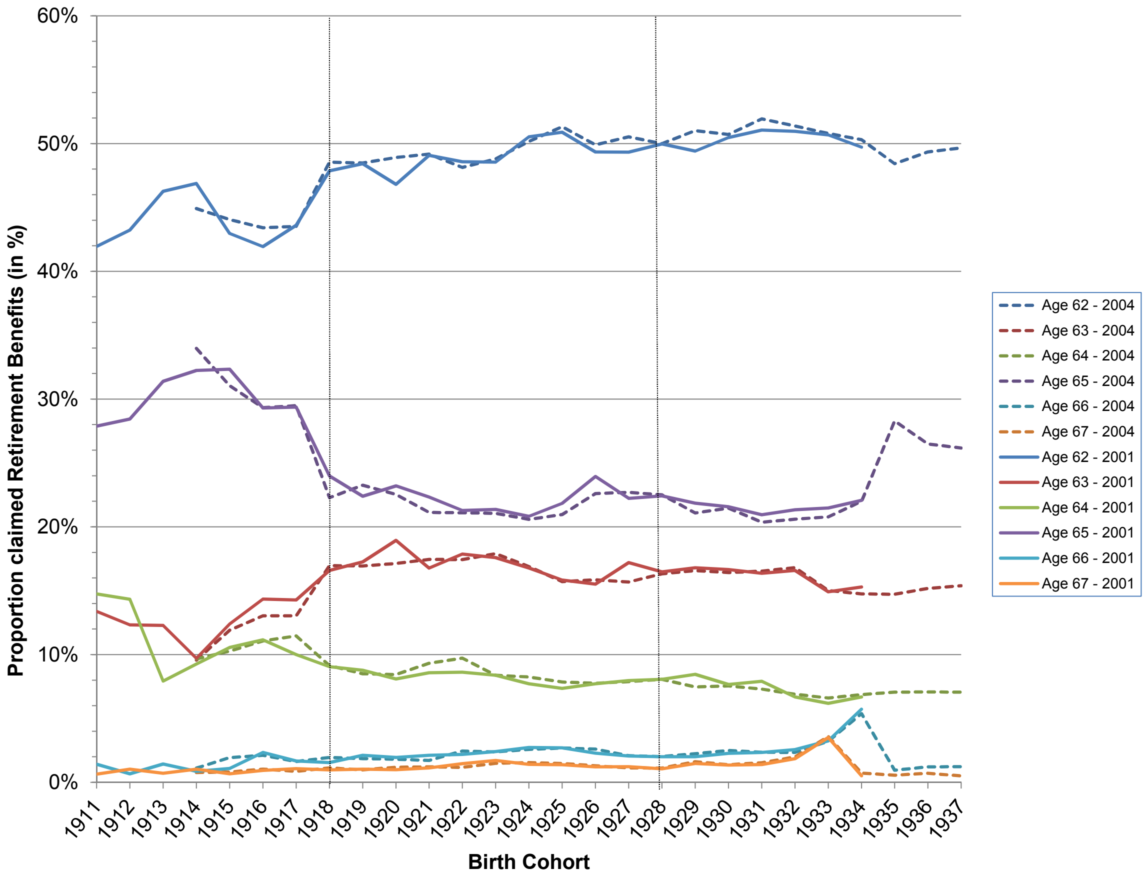

Figure 13: Distribution of Age at Retirement Benefit Take-up, Birth Cohorts 1911-1937 (Source: Authors' calculation based on 2001/2004 Public-Use Master Beneficiary Microdata)