

Retirement and Social Security Reform Expectations: A Solution to the New Early Retirement Puzzle[†]

Hugo Benítez-Silva[‡]
SUNY-Stony Brook

Debra S. Dwyer
SUNY-Stony Brook

Frank Heiland
Florida State University

Warren C. Sanderson
SUNY-Stony Brook

September 18, 2007

Abstract

Despite the persistent increase in life expectancy, combined with an increasing penalty for earlier benefit take-up under the Old Age Benefits Program of Social Security, people continue to claim benefits before the Normal Retirement Age (NRA) in record numbers. Furthermore, almost 60% of Americans claim at the earliest possible age. We solve and simulate a realistic, and empirically based dynamic life-cycle model of labor supply and benefit claiming, that accounts for the rich set of incentives affecting Older Americans. We model, among other sources of uncertainty, the uncertainty surrounding future benefit amounts, which can be rationalized by the perception of the need for reforms to the system. Using aggregate and individual level information on expected benefit cuts and probabilities of realizing them, this framework is one of the first to explain the large proportion of individuals claiming benefits early. Moreover, our model is the first to predict, consistent with the data, that even if the penalties for claiming benefits early increase (like with the ongoing increases in the NRA) the percentage of individuals claiming early might not necessarily decline.

Keywords: Retirement Benefits, Social Security Reform, Earnings Test, Dynamic Models.

JEL classification: J26, H55

[†] We are grateful to a number of employees of the Social Security Administration who have patiently answered our questions. Among those individuals we are especially thankful to Barbara Lingg, Christine Vance, Rona Blumenthal, and Joyce Manchester for their time, and for their insights into a number of the issues discussed in this paper. We are grateful to MRRC for the financial support through grant UM06-17. Benítez-Silva is also grateful for the financial support from NIH grant AG12985-02, the Spanish Ministry of Science and Technology through project number SEJ2005-08783-C04-01, and to the Fundación BBVA. Na Yin provided excellent research assistance.

[‡] **Corresponding Author:** Economics Department, SUNY-Stony Brook, Stony Brook, N.Y. 11794-4384, phone: (631) 632-7551, fax: (631) 632-7516, e-mail: hugo.benitez-silva@sunysb.edu

1 Introduction

Every American worker receives an annual Social Security Statement in which the government explicitly states that unless reforms are undertaken there will not be enough funds to pay benefits at the level at which they have been promised.¹ The need for reforms to Social Security does not seem to come as a surprise for the average American. In fact, as reported for example in several waves of the Health and Retirement Study, individuals believe that there is a 60% chance that in the next 10 years Social Security benefits will be reduced. Given that the Social Security Administration itself has decided to put the message in black and white for several years now, through the statements and in several reports that quantify the painful consequences of the lack of any kind of reform, makes the need of understanding the consequences of this reform uncertainty the more pressing. Researchers, however, have rarely modeled the uncertainty over Social Security reform and benefit levels, and little is known about how it affects the benefit claiming and retirement behavior of current retirees.

Public pensions are a major income source for older Americans, and under the Old Age and Survivor Insurance (OASI) system, the Social Security Administration paid benefits during 2006 to almost 41 million individuals who received about \$449.2 billion in benefits. Given its importance is not surprising that the discussion over the need of reforms to the system have gone on for a long time. Since the 1970s reforming Social Security has been a priority among economic researchers and policy makers. In fact, the 1983 Amendments were meant to solve the financial crisis that

¹ The statements make clear that by the time the Social Security Trust Fund is exhausted there will only be money to pay about 74% of the scheduled benefits. Similar statements can be found on the website of the Social Security Administration (<http://www.ssa.gov>). Even influential independent policy makers like former chairman of the Federal Reserve Alan Greenspan in front of the Committee of the Budget of the U.S. House of Representatives stated that *“Under current law, and even with the so-called Normal Retirement Age (NRA) for Social Security slated to move up to 67 over the next two decades, the ratio of the number of years that the typical worker will spend in retirement to the number of years he or she works will rise in the long term. A critical step forward would be to adjust the system so that this ratio stabilizes.”*

Social Security was headed for.² Within a decade of the passing of that legislation, it became clear that further reforms will be necessary to maintain the long-run balance of the system, given the evolution of the trust fund, and the trends in claiming behavior and labor supply of older Americans. Policy evaluation researchers failed to accurately predict the responses to this increase in the retirement age, casting doubt on the usefulness of the models that the analyses were based on. Specifically, ambiguous theoretical responses to such incentives were not properly analyzed, and what might have been considered large nominal cuts to benefits at the time were in fact much smaller reductions in real terms, especially given the large delays embedded in the implementation of the legislation.

The large retirement literature that developed during the 1980s and 1990s focused on explaining the connection between retirement incentives and retirement behavior.³ It concluded, quite convincingly, that the retirement peaks at age 62 and age 65 could be explained if the full set of incentives were included in the model. However, in the data used in those studies the majority of Americans were claiming benefits at age 65, while in the 1980s and 1990s the peak started to move towards age 62. By the end of the 1990s, almost 60% of older Americans were claiming benefits at age 62, and it has stayed at that level, even with the implementation of the 1983 Amendments that penalize early claiming of benefits and the substantial increase in expected longevity since the 1970s. In fact, as of July 2007, 70.7% of men and 75.6% of women claimed Social Security benefits before the Normal Retirement Age (NRA), compared to 36% and 59% in 1970, respectively.⁴ Clearly, the economic incentives seem to be insufficient to achieve the objective of prolonging

² For example, President Reagan, after signing the legislation stated “*Our elderly need no longer fear that the checks they depend on will be stopped or reduced. These amendments protect them. Americans of middle age need no longer worry whether their career-long investment will pay-off. These amendments guarantee it.*”

³ For a survey of this broad retirement literature see Lumsdaine and Mitchell (1999). Hurd (1990), Lumsdaine (1995), and Ruhm (1996) provide good discussions of the earlier literature.

⁴ See the Social Security Bulletin, OASDI Monthly Statistics, 1970 - 2007.

average work lives, given the strong correlation between benefit claiming and labor supply.

The purpose of this paper is to assess the extent to which the perceptions of future cuts might explain the puzzle of earlier take-up despite bigger penalties for doing so. If further cuts are anticipated, then individuals may be weighing early certain benefit amounts against expected reductions in the amount should they postpone retirement. Few papers have analyzed expectations over Social Security reforms. Büttler (1999) presents a fifteen periods Overlapping Generations model to analyze the effects of expected and unexpected reforms, and Phelan (1999) discusses more realistic characterizations of that model. One important difference between their work and ours is that we analyze individual-level behavioral effects of reform expectations that do not necessarily materialize in the short to medium run, instead of focusing on aggregate effects of reforms that do materialize. More recently, Sabelhaus and Topoleski (2007) analyze the aggregate effects of linking a future benefits level to economic and demographic outcomes. Interestingly, this link is what Phelan (2006) labels an *uncertainty minimizing policy*, but he finds that it cannot be sustained in equilibrium if the government type follows a Markov process, between opportunistic governments and trustworthy governments.

The present work shows that one of the keys to modeling the complex incentive structure of the US Old Age Benefits system is to account for the Earnings Test (ET) that directly impacts claiming and labor supply between the Early Retirement Age (ERA) and the NRA. Most researchers, have only focused on the taxation aspects of the Earnings Test provisions, and have not properly modeled the actuarial fairness of the system. Even the most sophisticated dynamic models of retirement have failed to explain the large proportion of Americans claiming early retirement without making assumptions about preferences for early retirement that are difficult to test and justify. We find that by accounting for the full set of incentives of the ET, and by modeling reform

expectations through the introduction of a small amount of uncertainty (based on self-reported responses to questions regarding expectations over future cuts in the Health and Retirement Study) of a benefit cut, we are able to match the claiming behavior observed in the data without relying on heterogeneous preferences as in Gustman and Steinmeier (2002). Those authors use a dynamic life cycle model that varies rate of time preference across individuals to explain the early claiming rates observed in the data. They measure rate of time preference using accumulated assets —under the assumption that those with higher rates would have saved less, driving their otherwise irrational desire to retire earlier. Notice that given the well known lack of non-parametric identification of dynamic structural models (see Rust 1994, Taber 2000, and Magnac and Thesmar 2002), the reliance on preference heterogeneity to explain behavior is less desirable than being able to account for it through the appropriate incentives, or even empirically grounded homogeneous beliefs about future events affecting economic constraints.

Any analysis of the effects of Social Security reforms should be performed within a model that can explain early take-up behavior and accounts for the full incentive structure. We find that a misspecified dynamic retirement model would erroneously predict that an increase in the NRA would delay claiming behavior and increase labor supply at older ages (see Gustman and Steinmeier 1985, for an early discussion of the possible consequences of the 1983 reforms). This kind of rationale seemed to have motivated (and still continues to motivate, rather surprisingly given the inadequate results of previous similar reforms) the advocates of reforms to the SSA system, who hoped to obtain further financial relief for the Social Security system from the additional tax revenue resulting from an extension of the work lives of later claimers. The policy debate around the time of the deliberations of the National Commission leading to the 1983 reforms of the system shows that the idea of affecting individuals' working-lifetime was at the forefront of the discussion

about the increases in the NRA (see Myers 1993, p. 316). This idea lived on, both in the popular press and in the policy arena, when those reforms were proven to be insufficient (see e.g., Brown 1996, p. 32, Rejda 1999, p. 112, World Bank 1994, pp. 323-324).

Once the appropriate ET incentives are modeled and the probability that the system will be reformed is accounted for, an increase in the NRA, consistent with the data, has little effect on claiming behavior and may even increase the proportion of individuals claiming before the NRA. This suggests that any policy that does not affect the ratio between the benefits received at different ages (for example, any further increases in the NRA, which effectively imply equal cuts of benefits across claiming ages) should not be expected to have much impact on claiming behavior. Therefore, the effect on labor supply is likely to be much weaker.

Section 2 provides additional background on retirement incentives and claiming behavior. Section 3 discusses Social Security reform expectations. Section 4 describes the dynamic structural model that we use to analyze the impact of uncertainty over Social Security reform, and the consequences of changes in the NRA. Section 5 presents the results of the simulations of the model, and discusses the policy implications of our findings. Section 6 concludes with a short discussion of policy alternatives.

2 The Old Age Benefits Incentive System

Social Security provides fairly complex incentives that affect the labor supply and benefit uptake behavior of individuals between the ERA and the maximum retirement age. These incentives are especially involved between the early and Normal Retirement Ages, and we analyze them in detail in the Appendix. Two of the most important incentives are the Social Security Earnings Test, which

determines the maximum level of earnings that do not result in a benefit reduction for individuals who have claimed retirement benefits before the NRA, and the Actuarial Reduction Factor (ARF), which determines the permanent reduction in benefits that individuals face if they claim benefits early. However, the role of the Earnings Test in the context of the adjustment of the ARF is not very well understood, or even known by many.

Although researchers have occasionally documented these fairly complex incentives, they have paid relatively little attention to the possible consequences of these provisions for labor supply and claiming behavior of early retirees.⁵ The existing research has primarily focused on the taxation aspects of the Earnings Test.⁶ Since the removal of the Earnings Test in the year 2000 for those above the NRA, there has been relatively little discussion of the Earnings Test for younger retirees, despite the fact that the arguments used against the former Earnings Test also apply to this case. In fact, the incentives provided by the Earnings Test for early retirees have remained essentially unchanged in the last three decades, but with a larger fraction of Americans retiring early, these incentives have become increasingly important. The literature has not addressed the implications of the possibility to affect the Actuarial Reduction Factor by working after claiming benefits and earning above the Earnings Test limit for labor supply and claiming behavior. We will show through our dynamic model that the appropriate modeling of these incentives is key in order to understand the claiming behavior of Older Americans.

Table 1, using data from Table 6.A4 of SSA's Statistical Supplement, shows how prevalent take-up at the earliest possible age has become. The peaks are at the eligibility ages of 62 and 65 which

⁵ Gustman and Steinmeier (1991), Myers (1993, p. 52), and Gruber and Orszag (2003) discuss this mechanism in some detail.

⁶ See Vroman (1985), Burtless and Moffitt (1985), Honig and Reimers (1989), Leonasio (1990), Reimers and Honig (1993), Reimers and Honig (1996), Friedberg (1998), Baker and Benjamin (1999), Friedberg (2000), and Votruba (2003).

comes as no surprise given this well established response to program incentives. Between 1994 and 2005, almost 60% of claimants have been taking their benefits at age 62, and between 15% and 20% wait for the normal age of retirement. A majority of the remaining individuals claim at age 63 or 64, with a very small proportion claiming after the NRA. The latter is worth emphasizing given that the Delayed Retirement Credit increased by half a percentage point every two years during this period. Notice the rather anomalous claiming behavior in 2000, which resulted in an increase in claiming at age 65, and a reduction of the proportion of individuals claiming at 62. This is driven by the large increase in new entitlements at age 65 and above in that year, very likely the product of the removal of the ET for those above the NRA, which made waiting to claim benefits because of a strong attachment to the labor force unnecessary. This conjecture is further supported by the evidence on benefits levels of the next table.

Table 2, also using data from the Statistical Supplement, shows the trends in benefits received, in dollars of 2005, as a function of the age at which benefits were claimed. We see a clear break in the patterns after 2000, especially in terms of the benefit levels at the NRA and above. In 1999 and 2000 later claiming led to consistently larger benefits, while the maximum benefit has been systematically obtained by those claiming at 65 since then. It drops sharply for those claiming after 65, potentially because those individuals are now of a type trying to catch up to compensate for a low wage career, or a sketchy one. Our interpretation of this evidence is that the removal of the ET for those above the NRA had the effect of allowing people to claim benefits independently of their labor supply behavior, leading relatively well-off individuals, who before waited to claim to avoid the ET, to claim sooner. Those claiming after the NRA are now either individuals trying to catch up after relatively lower wage career profiles, or spouses claiming on their partner's earnings histories. Notice that the scheduled increases in the NRA are essentially bringing back the old ET

for those above age 65, so the prediction is that a pre-ET-reform benefit level distribution is likely to emerge, at least in part, in the next years. It is important to emphasize that this table does not account for the actuarial reduction of benefits faced by individuals claiming before the NRA, or for the delayed retirement credit obtained by those after the NRA. In this research we are interested in the inflation-adjusted level of benefits actually received by claimers since this is what our dynamic model of retirement predicts.⁷

3 Social Security Reform Expectations

What the public debates on the need for Social Security Reform have certainly accomplished is to instill uncertainty over future Social Security benefit amounts, despite the increasing amount of information being provided regarding ones benefit amounts under current rules (Mastrobuoni, 2006). While it is true that individuals are better informed now than ever about how much they should expect to get under current rules given the start of regular mailings containing that information, people expect that these benefits are subject to change. And they know the direction of the change, one way or another benefits will be cut.

If further cuts are anticipated, then individuals may be weighing in early certain benefit amounts against expected reductions in the full amount should they postpone retirement. At the margin, we would expect people to continue to retire earlier despite the increasing penalties for doing so, if they expect the gain to waiting it out to be declining and possibly approaching zero.

To illustrate the mechanism we believe is at play regarding Social Security reform expectations, imagine two individuals with identical Primary Insurance Amounts (PIAs). One of them expects to

⁷ It is clear that analyzing the role of (theoretically) actuarially fair adjustments is important to understand the importance of individual heterogeneity in claiming behavior. Benítez-Silva and Yin (2007) focus on this point and find considerable individual heterogeneity in benefits receipt, especially for those above the NRA.

receive a given amount upon retirement at age 65 with certainty, while the other believes that there is a reasonable probability (could be quite small) of the benefit amount to decline for retirement at age 65 by the time he reaches that age. The incentives to hold out for the full benefit amount, at the margin, are higher for the person who believes the payoff is a certain amount. Hence, if individuals believe that holding out is risky, they might opt out earlier. This suggests that policies that cut benefits and reduce dependency ratios may not realize the desired effects, unless they are perceived as a complete solution to the financial crisis of the system. Otherwise, the concerns over solvency are likely to continue.⁸

To calculate the perceived probability that a drop in benefits will occur, we use data from a ten year panel based on the Health and Retirement Study (HRS). Specifically, we examine the responses to questions regarding expectations over future benefits between the years of 1992 and 2002. The responses reveal perceptions of a 60% probability, on average, of benefits becoming less generous within the next 10 years from individuals who began the survey at the pre-retirement eligibility ages of 51 to 61. The cumulative probability of that happening in a period of 10 years implies a 4.825% expectation of a drop in the next year. The HRS asked questions regarding the propensity for Social Security to make benefits less generous on a scale of 1 to 10. In the later years the question was altered slightly to add a time frame—some time in the next 10 years. This did not significantly alter the results. Given the magnitude of these responses, uncertainty over future benefits is warranted in our models of benefit take-up.

In order to gauge the expected size of the benefit cut, we note that a recent report by the Trustees of the Social Security system, SSA (2007), states that, in order to maintain solvency in

⁸ Other plausible explanations for the trends toward earlier claiming of benefits include increased longevity but worsened functional capacity for work (health as a taste shifter toward leisure), and an increased demand for leisure over time that offsets the higher price of leisure so that empirically it appears to be a Giffen good.

the long-run, the immediate and permanent cuts should be of around 13% of benefits in real terms. Individuals are likely responding to a more modest cut, predicting that the necessary burden will likely spread across cohorts. We assume individuals believe a permanent cut of 5.75% of benefits could occur, a level roughly equivalent to about a one year loss in benefits for the average person, and therefore a reasonable expectation for the average person.⁹

4 The Dynamic Model

The model used in this paper is closely related to those presented in Rust and Phelan (1997), and Benítez-Silva, Buchinsky, and Rust (2003 and 2006). Rust and Phelan (1997) did not model consumption and savings decisions, but did estimate the parameters of the model, using a Nested Fixed-Point algorithm, instead of calibrating them. Benítez-Silva, Buchinsky, and Rust (2003 and 2006) present the most closely related models, which are calibrated to match aggregate data and household level data from the Health and Retirement Study, and model the Social Security Disability Insurance decisions on top of the OASI incentives. Unlike the structural model developed in the present paper, these earlier models (or any other structural models we are aware of) do not explicitly account for the possibility of affecting the Actuarial Reduction, or the possibility of expecting a possible benefit cut in the future. Our model also shares a number of characteristics with the work of French (2005), van der Klaauw and Wolpin (2005), and Blau (2004) among other researchers who solve, simulate, and in some cases estimate, dynamic retirement models under uncertainty.

We assume that individuals live a maximum of 100 years, and face mortality probabilities sim-

⁹ The qualitative results we will present in the next sections are robust to the assumptions regarding probability of the drop, and the size of the drop. Quantitatively, however, they are robust as long as higher drops are expected with slightly smaller probabilities, and vice-versa.

ilar to those in the population. They start their working lives at age 21, and maximize the expected discounted stream of future utility, where the per period utility function $u(c, l, h, t)$ depends on consumption c , leisure l , health status h , and age t . We specify a utility function for which more consumption is better than less, with agents expressing a moderate level of risk aversion. The flip side of utility of leisure is the disutility of work. We assume that the utility (disutility of work) is an increasing function of age, is higher for individuals who are in worse health than individuals who are in good health, and is lower for individuals with higher human capital measured by the average wage. In addition, we assume that the worse an individual's health is, the lower their overall level of utility is, holding everything else constant. Moreover, we assume that individuals obtain utility from bequeathing wealth to heirs after they die. This model assumes that individuals are forward looking, and discount future periods at a constant rate β , assumed here to be equal to 0.96.

The model also allows for a variety of sources of uncertainty, like lifetime uncertainty, health uncertainty, wage uncertainty, and more importantly, Social Security benefits level uncertainty. We will see in the next section that the latter is essential to match the large peak of benefits claiming at age 62. Notice, however, that within the model, this uncertainty is never realized, and benefits are never cut, but the existence of a small probability of the event happening affects behavior, and results in claiming benefits earlier, consistently with the empirical evidence.

Any person who is not already receiving Social Security Old Age benefits is eligible to apply for OASI benefits.¹⁰ Individuals with at least 40 quarters of earnings covered for OASI before reaching their 62nd birthday are eligible to apply and benefit award is guaranteed. In the present version of the model we allow decisions to be made on an annual basis and assume no lag between

¹⁰ We are abstracting from Social Security Disability Insurance (SSDI), a program that allows workers with severe disabilities to receive Social Security benefits before the NRA. This program currently covers about 7 million Americans. See Benítez-Silva, Buchinsky, and Rust (2003 and 2006) for a life-cycle model of retirement and SSDI application.

application date and date of first receipt.

Calculation of benefits and the reduction factors are as explained in the Appendix on incentives for early retirement, assuming a NRA of 66. In particular the number of checks received in a year depends on the earnings after claiming: the number of checks (or the benefit amount on some checks received towards the end of the period) are reduced reflecting the 50% rate on labor incomes exceeding the Earnings Test limit between 62 and the January of the year a person turns 65 (33% thereafter). In other words, adjustments to benefits and ARFs occurs in accordance with the earnings and the Earnings Test limit, and we do not consider the possibility that beneficiaries ask Social Security for a reduction of benefits or return benefits received. Even though we set up an annual decision-making process, the Social Security Earnings Test is enforced semiannually, i.e. the benefits received by a beneficiary are adjusted, after reaching the NRA, for the earnings in excess of the Earnings Test limit, as long as six months or more, of benefits were withheld in the years between the early and Normal Retirement Ages. The structure and the details of the model are described below.

4.1 Model Details

We solve the dynamic life-cycle model by backward induction, and by discretizing the space for the continuous state variables.¹¹ The terminal age is 100 and the age when individuals are assumed to enter the labor force is 21. Prior to their 62nd birthday, agents in our model make a leisure and consumption decision in each period. At 62 and until age 70, individuals decide on leisure, consumption, and application for OASI benefits, denoted $\{l_t, c_t, ssd_t\}$, at the beginning of each period, where l_t denotes *leisure*, c_t denotes *consumption*, which is treated as a continuous decision

¹¹ See Rust (1996), and Judd (1998) for a survey of numerical methods in economics.

variable, and ssd_t denotes the individual's Social Security *benefit* claiming decisions.

After age 70 is assumed that all individuals have claimed benefits, and again only consumption and leisure choices are possible. Leisure time is normalized to 1, where $l_t = 1$ is defined as not working at all, $l_t = .543$ corresponds to full time work, and $l_t = .817$ denotes part-time work. These quantities correspond to the amount of waking time spent non-working, assuming that a full-time job requires 2000 hours per year a part-time job requires 800 hours per year. We assume two possible values for ssd_t . If ssd_t equals 1 the agent has initiated the receipt of benefits. If the individual has not filed for benefits or is not eligible then ssd_t is equal to 0.

If benefits are claimed before the NRA the monthly benefit amount is calculated similar to equation (9) in the Appendix. For a NRA of 66 years the reduction factor if claimed at 62 is 75%, 80.0% if claimed at 63, 86.67% if claimed at 63, and 93.33% if claimed at 65. Due to the Earnings Test, benefit initiation between the ERA and the NRA does not necessarily imply benefit receipt, nor is the reduction in the benefit rate necessarily permanent after the NRA as a result of the adjustment of the ARFs as discussed in the Appendix (see equation (10)). In particular, we use an annual Earnings Test limit of \$12,480 between 62 and 65 and \$33,240 between 65 and 66 (these numbers reflect the 2006 limits). In the former period benefits are reduced at a rate of \$1 per \$2 of earnings above the limit and \$1 per \$3 of earnings above the limit for the latter period. These are the correct rules for someone who turns 66 in December. Since those whose birthday is earlier in the year face the higher limit and lower tax rate for less than a year (January to month of birthday) we have also simulated two alternative versions, one with the \$12,480 limit throughout, and another using \$20,760, the midpoint between the two limits and a tax rate of 50%. The results of these models do not differ markedly from those presented in the paper and are available from the authors upon request. Those claiming after 66 earn the delayed retirement credit. We model

it following the rates faced by the 1943-1954 cohorts, of $\frac{2}{3}$ of 1% for each month not claimed between age 66 and 70.

We also incorporate a detailed model of taxation of other income, including the progressive federal income tax schedule (including the negative tax known as the EITC – Earned Income Tax Credit), and state and local income, sales and property taxes. Individuals whose combined income (including Social Security benefits) exceeds a given threshold must pay Federal income taxes on a portion of their Social Security benefits. We incorporate these rules in our model as well as the 15.75% Social Security payroll tax.

The model allows for four different sources of uncertainty: (a) lifetime uncertainty: modeled to match the Life Tables of the United States with age and health specific survival probabilities; (b) wage uncertainty: modeled to follow a log-normal distribution, function of average wages as explained in more detail below; (c) health uncertainty: assumed to evolve in a Markovian fashion using empirical transition probabilities from a variety of household surveys, including the NLSY79 and the HRS. The random draws to simulate these three sources of uncertainty are the same for all the models compared in this paper, such that the differences presented in the results are only due to the changes in the incentive schemes; (d) Social Security benefit level uncertainty: this is one of the main contributions to the paper and we explain it in detail below.

Regarding the latter type of uncertainty, we first assume that agents believe there is a chance of benefits being cut in the future. Second, at age 62 they believe that if they do not claim benefits then there is a small probability of those benefits being lower in the future. These beliefs are never realized in the simulations of the model, but are present in the expectations of the agents, resulting in possible changes in behavior. As explained in the previous section, we use reasonable parameter values based on aggregate data and household surveys.

The *state* of an individual at any point during the life cycle can be summarized by five state variables: (i) Current age t ; (ii) net (tangible) wealth w_t ; (iii) the individual's Social Security benefit claiming state ss_t ; (iv) the individual's health status, and (v) the individual's average wage, aw_t .¹² For computational simplicity, we assume that decisions are made annually rather than monthly, but we allow for the benefit adjustments due to earnings above the Earnings Test limit to happen semi-annually. This means that although individuals can only decide to claim benefits at the time they turn 62, 63, etc. their Social Security state can be updated every year, depending in their labor earnings, to reflect that their benefits will be adjusted for benefits withheld for periods of six months, or one year. Since the adjustment in benefits becomes effective only after they reach the NRA individuals still receive benefits at the original claiming rate in the period between the time of withholding of benefits until the NRA, consistent with current rules.

The ss_t variable can assume up to fourteen mutually exclusive values between 62 and 66: $ss_t = 0$ (not entitled to benefits), $ss_t = 62$ (entitled to OASI benefits at the ERA), and $ss_t = 62.5, 63, 63n, 63.5, 64, 64n, \dots, 65.5, 66, 66n$ represents the remaining 12 Social Security states corresponding to the level of benefits individuals will receive when they reach the NRA. For individuals who decide to claim after the NRA, ss_t can take four additional values, age 67 to 70, since everyone is assumed to claim no later than age 70. We created an additional (implicit state) variable, ssn_t , which can assume up to five mutually exclusive values: $ssn_t = 0$ (all benefits received, i.e. no benefits withheld), $ssn_t = 1$ (representing an original claim at age 62 of someone who had some benefits withheld; this applies, for example, to individuals with a ss_t equal to 62.5, 63n,

¹² This translates into a problem with over half a million states in which to solve the model (80 periods, 15 discretized wealth states, 8 discretized average wage states, 3 health states, and 18 Social Security states). We are able to solve this model and simulate it 10,000 times in under 20 minutes in a Dual-Processor Linux Machine with 3.6GHz Xeon Processors using Gauss, and exploiting its capability to link dynamic libraries written in C by the authors and some of their co-authors. These C libraries perform over 95% of the computations involved in solving and simulating these models. The code used for these simulations is available upon request, and will eventually be available on the web.

or $64n$), $ssn_t = 2$ (representing an original claim at age 63 for someone who had some benefits withheld), $ssn_t = 3$ (representing an original claim at age 64 for someone who had some benefits withheld), etc. With this structure we are able to separate, for example, whether someone is a 63 claimer, denoted by $ss_t = 63$, or is really a 62 claimer who has accumulated one year of withheld benefits, represented here by $ss_t = 63n$. These two individuals will receive the same amount of benefits after the NRA, but their benefit would differ before the NRA, as explained in the Appendix, and in additional detail in Benítez-Silva and Heiland (2006, and 2007).

In addition to age, wealth, health, Social Security status, Benefit Adjustment status, and current income, the average indexed wage is a key variable in the dynamic model, serving two roles: (1) it acts as a measure of *permanent income* that serves as a convenient *sufficient statistic* for capturing serial correlation and predicting the evolution of annual wage earnings; and (2) it is key to accurately model the rules governing payment of the Social Security benefits. An individual's highest 35 years of earnings are averaged and the resulting *Average Indexed Earnings* (AIE) is denoted as aw_t . The PIA is the potential Social Security benefit rate for retiring at the NRA. It is a piece-wise linear, concave function of aw_t , whose value is denoted by $pia(aw_t)$.

In principle, one needs to keep as state variables the entire past earnings history. To avoid this, we follow Benítez-Silva, Buchinsky, and Rust (2006) and approximate the evolution of average wages in a Markovian fashion, i.e., period $t + 1$ average wage, aw_{t+1} , is predicted using only age, t , current average wage, aw_t , and current period earnings, y_t . Within a log-normal regression model, we follow Benítez-Silva, Buchinsky, and Rust (2003), such that:

$$\log(aw_{t+1}) = \gamma_1 + \gamma_2 \log(y_t) + \gamma_3 \log(aw_t) + \gamma_4 t + \gamma_5 t^2 + \varepsilon_t. \quad (1)$$

The R^2 for this type of regression is very high, with an extremely small estimated standard error, resulting from the low variability of the $\{aw_t\}$ sequences. This is a key aspect of the model given

the important computational simplification that allows us to accurately model the Social Security rules in our DP model with minimal number of state variables.

We then use the observed sequence of average wages as regressors to estimate the following log-normal regression model of an individual's annual earnings:

$$\log(y_{t+1}) = \alpha_1 + \alpha_2 \log(aw_t) + \alpha_3 t + \alpha_4 t^2 + \eta_t. \quad (2)$$

This equation describes the evolution of earnings for full-time employment. Part-time workers are assumed to earn a pro-rata share of the full-time earnings level (i.e., part-time earnings are $0.8 \cdot 800/2000$ of the full-time wage level given in equation (2)). The factor of 0.8 incorporates the assumption that the rate of pay working part-time is 80% of the full-time rate. Using the history of earnings from the restricted HRS data set we obtained very high R^2 using this methodology.

The advantage of using aw_t instead of the actual Average Indexed Earnings is that aw_t becomes a sufficient statistic for the person's earnings history. Thus we need only keep track of aw_t , and update it recursively using the latest earnings according to (1), rather than having to keep track of the entire earnings history in order to determine the 35 highest earnings years, which the AIE requires.

For the 1943-1954 cohort the NRA is 66 and the PIA is permanently reduced after the NRA by an actuarial reduction factor of $\exp(-g_1(k - adjm))$, where k is the number of years prior to the NRA but after the ERA that the individual first starts receiving OASI benefits and $adjm$ corrects for periods where no benefits were received due to earnings above the Earnings test limit. Before the NRA, benefits are reduced by an actuarial reduction factor of $\exp(-g_1 k)$. In the absence of adjustments to the ARFs, the actuarial reduction rate for the 1943 to 1954 cohort is $g_1 = .0713$, which results in a reduced benefit of 75% of the PIA for an individual who first starts receiving OASI benefits at age 62 in the absence of any adjustments of the ARFs. In the policy simulations

that increase the NRA to 67, the reduced benefit at age 62 is 70% of the PIA.

To increase the incentives to delay retirement, the 1983 Social Security reforms gradually increased the NRA from 65 to 67 and increased the delayed retirement credit. This is a permanent increase in the PIA by a factor of $\exp\{g_2 l\}$, where l denotes the number of years after the NRA that the individual delays receiving OASI benefits. The rate g_2 is being gradually increased over time. The relevant value for the 1943 to 1954 cohort is $g_2 = 0.0769$, which corresponds to an increase in 8% in benefits per year of delay after the NRA. The maximum value of l is $MRA - NRA$, where MRA denotes a “maximum retirement age” (currently 70), beyond which further delays in retirement yield no further increases in PIA. Clearly, it is not optimal to delay applying for OASI benefits beyond the MRA , because due to mortality, further delays generally reduce the present value of OASI benefits the person will collect over their remaining lifetime.

We assume that the individual’s utility is given by

$$u_t(c, l, h, age) = \frac{c^\gamma - 1}{\gamma} + \phi(age, h, aw) \log(l) - 2h, \quad (3)$$

where h denotes the health status and $\phi(age, h, aw)$ is a weight that can be interpreted as the *relative disutility of work*. We use the same specification for ϕ and the disutility from working as in Benítez-Silva, Buchinsky, and Rust (2006). The disutility of work increases with age, and is uniformly higher the worse one’s health is. If an individual is in good health, the disutility of work increases much more gradually with age compared to the poor health, or disabled health, states. The disutility of work decreases with average wage. We postulate that high wage workers, especially highly educated professionals, have better working conditions than most lower wage blue collar workers, whose jobs are more likely to involve less pleasant, more repetitive, working conditions and a higher level of physical labor.

We assume that there are no time or financial costs involved in applying for OASI benefits. The

parameter γ indexes the individual's level of risk aversion. As $\gamma \rightarrow 0$ the utility of consumption approaches $\log(c)$. We use $\gamma = -.37$, which corresponds to a moderate degree of risk aversion, i.e., implied behavior that is slightly more risk averse than that implied by logarithmic preferences.

Let $V_t(w, aw, ss, h)$ denote the individual's value function, the expected present discounted value of utility from age t onward for an individual with current wealth w , average wage aw , in Social Security state ss and health state h . We solved the DP problem via numerical computation of the Bellman recursion for V_t given by

$$V_t(w, aw, ss, h) = \max_{\substack{0 \leq c \leq w \\ l \in \{.54, .81, 1\} \\ ssd \in A_t(ss)}} V_t(w, aw, ss, c, l, ssd, h), \quad \text{where} \quad (4)$$

$$V_t(w, aw, ss, c, l, ssd, h) = u_t(c, l, h) + \beta[1 - d_t(h)]EV_{t+1}(w, aw, ss, c, l, ssd, h), \\ + d_t(h)EB(w, aw, ss, c, l, ssd, h). \quad (5)$$

where $A_t(ss)$ denotes the set of feasible Social Security choices for a person of age t in Social Security state ss and $d_t(h)$ denotes the age and health-specific mortality rate, $B(w)$ is the bequest function, and EB denotes its conditional expectation. We have used the HRS and AHEAD data to estimate age and health-specific death rates, but since there is little data on individuals over 80 years old we make parametric smoothness assumptions on the $d_t(h)$ function (basically a logit functional form that is polynomial in t and has dummy variables for the various health states h) and subject the estimates to the further restriction that for each t the expected hazard over h should equal the unconditional age-specific death rates given in the 1997 edition of the U.S. Decennial life tables.¹³ The function EV_{t+1} denotes the conditional expectation of next period's value function, given the individual's current state (w, aw, ss, h) and decision (c, l, ssd) . Specifically, we have

¹³ De Nardi, French, and Jones (2006) find that more sophisticated mortality characterizations do not seem to significantly improve the fit of a related dynamic structural model which focuses on post-retirement saving behavior.

$$EV_{t+1}(w, aw, ss, c, l, ssd, h) = \int_{y'} \sum_{h'=0}^2 \sum_{ss'=0}^{17} V_{t+1}(wp_t(w, aw, y', ss, ssd), awp_t(aw, y'), ss') \times f_t(y'|aw)k_t(h'|h)g_t(ss'|aw, w, ss, ssd)dy', \quad (6)$$

where $awp_t(aw, y)$ is the Markovian updating rule that approximates Social Security's exact formula for updating an individual's average wage, and wp_t summarizes the law of motion for next period's wealth, that is,

$$wp_t(w, aw, y, ss, ssd) = R [w + ssb_t(aw, y', ss, ssd) + y' - \tau(y', w) - c], \quad (7)$$

where R is the return on saving, and $\tau(y, w)$ is the *tax function*, which includes income taxes such as Federal income taxes and Social Security taxes and potentially other types of state/local income and property/wealth taxes. The awp_t function, derived from (1), is given by

$$awp_t(aw, y) = \exp \{ \gamma_1 + \gamma_2 \log(y) + \gamma_3 \log(aw) + \gamma_4 t + \gamma_5 t^2 + \sigma^2 / 2 \}, \quad (8)$$

where σ is the estimated standard error in the regression (1). Note there is a potential ‘‘Jensen's inequality’’ problem here due to the fact that we have substituted the conditional expectation of w_{t+1} into the next period value function V_{t+1} over w_{t+1} and aw_{t+1} jointly. However, as noted above, the R^2 for the regression of aw_{t+1} on aw_t is virtually 1 with an extremely small estimated standard error $\hat{\sigma}$. In this case there is virtually no error resulting from substituting what is an essentially deterministic mapping determining aw_{t+1} from w_{t+1} and aw_t .

Above, $f_t(y|aw)$ is a log-normal distribution of current earnings, given current age t and average wealth aw , that is implied by (2) under the additional assumption of normality of errors η_t . The discrete conditional probability distributions $g_t(ss'|aw, w, ss, ssd)$ and $k_t(h'|h)$ reflect the transition probabilities in the Social Security and health states, respectively.

Finally, to account for the Social Security reform expectations, we introduce in the optimization process the possibility that with a probability of 4.825%, the agents faced a modified equation (4), with a modified equation (6) and modified law of motion for wealth, which embeds a permanent drop in benefits of 5.75%.

5 Simulation Results and Policy Experiments

Table 3 reports results from three different models of Social Security, assuming a NRA of 66.¹⁴ Model 1 treats the Earnings Test as a pure tax on earnings above the corresponding ET limit, which is how it may be perceived by a proportion of the general public given the difficulties in understanding the role of the adjustment factors when working beyond benefit take-up (Benítez-Silva and Heiland, 2006 and 2007), and how a majority of researchers have modeled these incentives. Using this framework, our model of optimal behavior predicts that only about 35% of claimers would take up at age 62, with a much smaller peak at 65 at roughly 18%, with a bulk of the remaining beneficiaries claiming at the ages in between. Benefits, given this behavior, increase slightly with age at all points so that there are economic incentives for delaying take-up.

With the implementation of the proper Earnings Test incentives, which allow for the modification of the actuarial reduction factor through work after take-up with earnings above the ET limit, Model 2 shows a trend towards earlier claiming of benefits. This moves us closer to the actual take-up rates with a jump from only 35% claiming at 62 in Model 1 to over 48% in Model 2. The actual current take-up rate is around 56.6% (SSA Statistical Supplement, 2006). There continues to be a second smaller peak at age 65 with 23% of the sample holding out for the unpenalized NRA

¹⁴ In this paper we focus on claiming behavior and labor supply, but the model also simulates the evolution of wealth, consumption, and wages over the life cycle. As shown in Benítez-Silva, Buchinsky, and Rust (2006), in a related model, the predictions of the model are consistent with the HRS data.

benefit. Only 4% of the sample hold off their benefit take-up until after 65 when the penalty for working after take-up is reduced, compared to 11% in Model 1.

While the biggest gains towards actual behavior come from adding the possibility of affecting the adjustment factors to the model, the predictions match the actual pattern closely once Social Security benefit uncertainty is added (Model 3). The predicted benefit take-up when individuals perceive a potential risk to delaying take-up increases to 59%, compared to the 57% actual take up rates at this age. The peak at age 65 falls to 21% compared to predictions from Model 2 (which predicts 23% take-up), and this again is closer to the actual rate of 19%. More people are willing to take lower benefits earlier, knowing they can move closer to the full benefit with the ARF adjustments, and preferring the actuarial reduction with certainty over future uncertain cuts.

The comparison between Models 2 and 3 provides a sense of the effects of reducing uncertainty over reforms. We can see that Model 2 predicts more full-time work, especially at age 62, but also in later ages, due to the strong connection between claiming behavior and labor supply. These results are broadly consistent with the findings of Büttler's (1999), who indicates that uncertainty considerations are important, and that governments can reduce the amount of uncertainty agents have to deal with, resulting in welfare improving allocations, by providing better information regarding the timing and the type of reforms in store.

One of the advantages of dynamic models is that it allows us to perform a welfare analysis. We have computed compensating variations, which capture the willingness to pay, or in this case the need to be compensated, for having to face the uncertainty over Social Security reforms. We find that given that the uncertainties are never realized, the differences in welfare are very small, and affect only a small proportion of individuals. Only about 5% of individuals in our simulations see a drop in their welfare, and among those the drops in welfare account for less than 1.5% of

their average wealth in the simulations. This suggests that even though the changes in behavior resulting from this source of uncertainty are clearly non-trivial, many of the individuals forced to claim earlier were originally close to indifferent with respect to claiming at other ages.

The last column in the three panels of Table 3 shows the benefit levels predicted by the model. Notice how close they are to the actual benefits received by Older Americans, which we reported in Table 2. This provides further confirmation that our model accurately matches the observed benefit distribution qualitatively, and quantitatively, even in dollar terms. Given that we are using a NRA of 66, and individuals face the Earnings Test between age 65 and 66 as a result, the relationship between benefit levels at different ages is closer to that present in the period before the elimination of the ET for those above the NRA. This translates in the prediction that later claimers obtain higher benefits.

Table 4 simulates behavioral responses to an increase in the NRA to 67. We resolve and re-simulate the same three models with this modification. If the Earnings Test only had the taxation aspect, an increase in the NRA to 67 would yield the intended behavioral responses, delayed benefit take-up and greater incidence of retirement. With the original NRA of 66, more than half of the beneficiaries claim by age 63. When the NRA increases about half of the beneficiaries are claiming at ages 65 and 66. As shown in the last column, this group is working more to earn the same benefits, consistent with the expected effects of a cut in benefits. Overall, the predicted responses do not follow the patterns in the data very closely.

Adding in the appropriate ET rules leads us to conclude that with a more realistic set of incentives, the predictions move significantly closer to the observed behavior. With the higher NRA, the take up rate at 62 increases holding the rest of the features of Model 2 constant. There is no longer a peak at 65 nor is there one at 67. The next peak after age 62 is at age 66. Model 3, which allows

for uncertainties over future benefit guarantees, captures the second peak at age 65, consistent with the pattern observed in the data. The results suggest that increases in the NRA might not yield the expected behavioral consequences; in fact, they may result in even earlier claiming of retirement benefits.

These findings support the sound economic intuition that any policy that does not affect the ratio between the benefits received at different ages (for example, any further increases in the NRA, which cut benefits equally across claiming ages) should not be expected to have much impact on claiming behavior, and therefore relatively little is to be expected in terms of labor supply responses, except maybe some increases in order to compensate for the loss of benefits.

6 Conclusions

We develop a dynamic life cycle framework that accounts for the majority of the complexities of the Old Age Social Security program, as well as expectations of future benefit cuts, to predict benefit claiming behavior and labor supply. The attention to the details of the incentives provided by OASS and the modeling of plausible expectations regarding future benefit cuts pays off, as we are able to explain the puzzle of the large proportion of Older Americans claiming benefits early despite rising longevity. We use the framework to show that the scheduled increases in the Normal Retirement Age, and possible further increases that are being debated, have little impact on the retirement incentives when properly modeled. This casts doubt on the efficacy of any reforms that cut benefits equally across claiming ages to promote lasting changes in behavior. From a policy perspective, our simulation results illustrate the importance of basing policy analysis on models that account for the actual incentive schemes and more realistic beliefs about future benefits. Predictions of

overly simplistic or clearly misspecified models can easily be erroneous and intuitions based on simplistic models can lead to large miscalculations.

Given the reality of the Social Security burden, cuts will have to be made one way or another. Our results suggest that mandating a set of cuts, with a serious commitment to resolving the solvency issues of the program for more than a couple of decades, so that no further cuts are expected, should delay take-up and alleviate, at least in part, the financial crisis. However, the necessary cuts for this plan to succeed might impose undue burden on a generation, potentially resulting in severe welfare losses. The challenge can be understood in the context of the work of Auerbach and Hassett (2006), who discuss sticky policies. Unfortunately when it comes to Social Security, policies are not sticky enough so that there is little confidence in future benefit amounts among planners.

Recent work by Phelan (2006) shows that in this kind of situations the unique subgame perfect equilibrium—when government type follows a Markov process—for this case would involve periodic benefits cuts followed by long periods of moderate taxation to restore the trust on the system by individuals. Interestingly, the government (and the individuals) would be better off if it could credibly commit to a time invariant policy. It appears that with respect to its Social Security system, the U.S. may be in such an equilibrium. Given that federal officials frequently stress that painful reforms are still ahead, the population has yet to regain confidence in the future of the system.

Given that the benefit cuts necessary to make the system solvent in the long term (such as immediate and permanent cuts of around 13% of benefits, equivalent to increasing the NRA to age 68 as of 2008, or an immediate and permanent increase in the payroll tax of 1.95 percentage points, as discussed by the Trustees of the Social Social Security Administration, SSA (2007, p. 3)) are likely to be politically unfeasible, it is natural to consider less painful reforms such as the elimination of the Earnings Test, the increase in the Early Retirement Age, or the adoption of a benefit schedule

that rewards later claimers. These reforms are unlikely to change individuals' expectations about the solvency problems of the system, but may affect claiming behavior and labor supply. The elimination of the ET might not have a large enough effect on labor supply to meaningfully address the solvency problems of the system and likely leads to even earlier claiming.¹⁵ The increase in the Early Retirement Age would mechanically delay claiming benefits and likely increase labor supply in the years leading to the new ERA, but would have a limited effect on the long run solvency of the system. Furthermore, it is highly regressive given the socio-demographic composition of early claimers. An alternative to those policies is to devise an actuarially unfair incentive structure in which late claiming is rewarded via higher benefits. The latter would be implemented hoping that the actual labor supply responses, and the resulting tax revenues, compensate for the cost of such a policy. While it is regressive, it might be less so than just increasing the Early Retirement Age.

¹⁵ See Benítez-Silva and Heiland (2007), Song and Manchester (2007), French (2005), Song (2004), Gustman and Steinmeier (2004), Gruber and Orszag (2003), and Disney and Smith (2001) for a discussion of the labor supply effects of the removal of the Earnings Test.

Appendix: Social Security Incentives for Early Retirement

Individuals who claim benefits before the NRA but continue to work or reenter the labor force can reduce the early retirement penalty by suspending benefit payments.¹⁶ The Actuarial Reduction Factor (or early retirement reduction factor), in turn, will be increased proportionally to the number of months without benefits, which will increase benefits permanently after the individual reaches the NRA.¹⁷ This adjustment of the ARF allows those who become beneficiaries before the NRA to partially or completely reverse the financial consequences of their decision, averting being locked-in at the reduced rate. In the sequel of this section the exact details of these incentives are presented.

Benefit Calculation

Individuals aged 62 or older who had earned income that was subject to the Social Security payroll tax for at least 10 years since 1951 are eligible for retirement benefits under the Old Age benefits program. Earnings are subject to the tax up to an income maximum that is updated annually according to increases in the average wage.¹⁸ To determine the monthly benefit amount (MBA), the Social Security Administration calculates the Primary Insurance Amount (PIA) of a worker as a concave piece-wise linear function of the worker's average earnings subject to Social Security taxes taken over her 35 years of highest earnings. If the benefits are claimed at the NRA (66 for those born between 1943 and 1954, and currently at 65 and 8 months), the MBA equals the PIA. If an individual decides to begin receiving benefits before the NRA and exits the labor force or stays below the earnings limit, her MBA is reduced by up to 25%, assuming a NRA of 66. Under the current regulation of the OA program, the monthly benefit amount received upon first claiming benefits depends on the age (month) of initiation of Social Security benefits, in the following way,

(9)

$$MBA_t = \begin{cases} (0.75 + 0.05 * \frac{1}{12} * (\text{Months not claimed in the period prior to 3 years before NRA})) * PIA & \text{if claimed more than 3 years before NRA;} \\ (0.80 + 0.20 * \frac{1}{36} * (\text{Months not claimed in 3 years before NRA})) * PIA & \text{if claimed within the 3 years before NRA.} \end{cases}$$

where MBA_t represents the monthly benefit amount before the NRA (see SSA-S 2005, p.18). Assuming that the individual continues to receive benefits, her MBA_t is permanently reduced. The Actuarial Reduction Factor (ARF) underlying this calculation is a permanent reduction of benefits by 5/9 of 1 percent per month for each month in which benefits are received in the three years

¹⁶ In this paper, we are not considering spousal benefits and joint decision making in the household. The complexities introduced by those considerations are out of the scope of this analysis. See Gustman and Steinmeier (1991), Coile, Diamond, Gruber, and Josten (2002), and Votruba (2003) for a discussion. By ignoring spousal benefits we are not taking into account the fact that approximately 5.96% 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 Public-Use Microdata File provided by the Social Security Administration and refers to a 1% random sample of all beneficiaries as of December of 2001.

¹⁷ Given a NRA of 66, which will be the prevailing one for the cohort born between 1943 and 1954, the Actuarial Reduction Factor is a number between 0.75 and 1 depending on when the individual claims benefits, and how many months he or she earns above the Earnings Test after claiming benefits.

¹⁸ Six percent of the 153 million workers with Social Security taxable earnings in 2002 had earnings at or above the maximum amount.

immediately prior to the NRA. The reduction of benefits is 5/12 of 1 percent for every month before that. Thus, the maximum actuarial reduction will reach 30 percent as the NRA increases to 67 over the next few years (see SSA-S 2005, p.18).¹⁹

Actuarial Reduction Factor

One less-emphasized feature of the process of benefit reduction due to early retirement is the possibility to reduce the penalty even after initiating the receipt of benefits. The specifics of this adjustment to the Actuarial Reduction Factor are documented in the Social Security Handbook (SSA-H, §724. *Basic reduction formulas*, §728. *Adjustment of reduction factor at FRA*) and in the internal operating manual used by Social Security field employees when processing claims for Social Security benefits (SSA-M, RS00615. *Computation of Monthly Benefits Amounts*) but may not be well-understood by the retirees.²⁰ To illustrate this feature of the system, suppose the NRA is 66 years, and an individual claims benefits at age 62 and n months, where $n \ll 48$, receives checks for x months where ($n + x \ll 48$), and suspends receiving checks after that until she turns 66 (after which she retires for good). In this case she receives x checks of

(10)

$$MBA_t = \begin{cases} (0.75 + 0.05 * \frac{1}{12} * n) * PIA & \text{if claimed more than 3 years before NRA;} \\ (0.80 + 0.20 * \frac{1}{36} * n) * PIA & \text{if claimed within the 3 years before NRA.} \end{cases}$$

After turning 66, her MBA will be permanently increased to

$$MBA_t = [0.75 + (0.20 * \frac{1}{36} * n) + (0.20 * \frac{1}{36} * (36 - n - x)) + 0.05] * PIA. \quad (11)$$

It is important to note that the adjustment of the ARF is automatic and becomes effective only after reaching the NRA.

Earnings Test

The Earnings Test limit defines the maximum amount of income from work that a beneficiary who claims benefits before the NRA under OASI may earn while still receiving the “full” MBA .²¹

¹⁹ The reductions in benefits for early claimers are designed to be approximately actuarially fair for the average individual. During the post-NRA period additional adjustments exist: 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 percent for each month up to age 70 which is considered actuarially fair. For those born before 1943 it ranges from 11/24 to 5/8 of 1 percent per month, depending on their birth year.

²⁰ The Social Security Administration does not use the term Actuarial Reduction Factor in their publications, but a number of the people we have talked to within the administration do use this terminology. In publications the related concept of “Reduction Factor(s)” (RF) which is simply the number of months in which benefits were received before the NRA is used. The RF maps into a “Fraction” that ranges between 0.75 and 1 (for an ERA of 62 and an NRA of 66). The latter corresponds to what we refer to as ARF. The ARF (“Fraction”) is adjusted upwards at the NRA according to the number of months before the NRA in which benefits were withheld.

²¹ Some sources of income do not count under the Earnings Test. For details see SSA-H §1812. Notice that retirement contributions by the employer do not count towards the limit, but additional contributions by the employee even if they are through a payroll deduction are counted. This means that individuals earning above the limit cannot just increase their retirement savings to avoid being subject to the limit. We thank Barbara Lingg and Christine Vance from the Social Security Administration for clarifying this point, which is rarely discussed in any publication.

Earnings above the limit are taxed at a rate of 50 percent for beneficiaries between age 62 and the January of the year in which they reach the NRA, and 33 percent from January of that year until the month they reach the NRA (SSA-S 2005, p.19; SSA-S 2005, Table 2.A18). For the latter period, the earnings limit is higher, \$31,800, compared with \$12,000 for the earlier period as of 2005 (SSA-S 2005, Table 2.A29). Starting in 2000, the Earnings Test was eliminated for individuals over the NRA.

Individuals who continue or reenter employment after claiming Social Security benefits before the NRA, and whose earning power or hours constraints are such that their income from work is around or below the earnings limit, are mailed their full monthly check from Social Security and are locked-in at the reduced benefit rate permanently. Those with earnings above the limit will not receive checks from Social Security for some months and thereby adjust their ARF.²² Individuals have the option of informing Social Security to suspend the monthly benefit payment at any time if they believe they will be making earnings high enough above the Earnings Test. However, during the first year after claiming benefits, the Social Security Administration performs a monthly test to determine whether the person should receive the monthly check. As a result an early claimer who is not working or earns below the limit in the months after claiming (“grace year”) will receive all monthly benefits even if earnings for that calendar year exceed the Earnings Test limit due to high earnings before claiming.²³ After the first year, the test is typically yearly and it depends on the expected earnings of the individual. Given the scarce documentation of the functioning of the ARF, having earned above the earnings limit, and thus receiving fewer checks, may be a common way for beneficiaries to learn about the possibility of undoing the early retirement penalty.²⁴

²² A beneficiary may receive a partial monthly benefit at the end of the tax year if there are excess earnings that do not completely offset the monthly benefit amount (see SSA-H, §1806).

²³ Social Security claim specialists emphasized to us that during the first year after claiming they do what is most advantageous to the claimer, the monthly or the yearly test, if they have enough information. However, they failed to clarify what that means. Some of them said the number of checks individuals receive is maximized, but we were unable to find documentation of such practices. In any case, the internal operating instructions used by Social Security field employees when processing claims for Social Security benefits state that the monthly Earnings Test only applies for the calendar year when benefits are initiated unless the type of benefit changes (see SSA-M, RS02501.030).

²⁴ See Benítez-Silva and Heiland (2006) for a numeric example of the streams of income resulting from these incentives.

Table 1: 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
<i># of Claimants^b</i>	<i>1,444.5</i>	<i>1,424.8</i>	<i>1,396.1</i>	<i>1,418.9</i>	<i>1,441.3</i>	<i>1,484.6</i>	<i>1,758.9</i>	<i>1,574.0</i>	<i>1,595.5</i>	<i>1,593.3</i>	<i>1,680.3</i>	<i>1,793.5</i>

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.

Table 2: Social Security beneficiaries' monthly benefits by age, 1994-2005. In Dollars of 2005.

Age/Year	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Age 62	788.58	789.53	785.31	802.01	815.35	855.64	864.56	884.42	892.58	900.40	888.31	881.9
Age 63	882.14	906.02	942.89	881.71	907.85	928.79	960.51	973.08	1,002.77	1,006.43	996.66	986.9
Age 64	981.51	982.52	997.16	1,014.71	1,001.66	987.87	1,020.39	1,072.88	1,119.80	1,119.68	1,102.01	1,089.8
Age 65	1,083.9	1,091.07	1,087.78	1,117.29	1,088.05	1,100.29	1,184.50	1,176.10	1,239.22	1,257.03	1,270.85	1,298.3
Age 66	1,022.36	1,077.35	1,033.01	1,049.14	1,030.42	1,093.73	1,247.67	939.56	881.73	919.08	981.26	1,052.2
Age 67	1,027.76	1,138.1	1,071.35	988.67	1,050.19	1,128.66	1,285.78	911.44	873.48	877.89	933.59	1,010.4

Table 3: Simulation Results: 10,000 Simulations of the Dynamic Model

Ages	Survivors	Full-Time^a	Part-Time^a	No Work^a	Claimers^b	Benefits in \$
Model 1: Earnings Test as a Tax						
Age 60	8,234	5,749 (69.8%)	163 (5.7%)	2,322 (28.8%)	—	—
Age 61	8,078	5,635 (69.7%)	213 (2.6%)	2,230 (27.6%)	—	—
Age 62	7,951	4,714 (59.2%)	2 (0.02%)	3,235 (40.6%)	2,672 (34.9%)	1,042
Age 63	7,762	2,013 (25.9%)	856 (11.0%)	4,893 (63.0%)	1,331 (17.4%)	1,151
Age 64	7,586	495 (6.5%)	2,008 (26.4%)	5,083 (67.0%)	1,048 (13.7%)	1,272
Age 65	7,420	113 (1.5%)	2,731 (36.8%)	4,576 (61.6%)	1,362 (17.8%)	1,391
Age 66	7,239	414 (5.7%)	3,484 (48.1%)	3,341 (46.1%)	847 (11.0%)	1,500
Model 2: Earnings Test with ARF Adjustments						
Age 60	8,234	5,749 (69.8%)	154 (1.8%)	2,331 (28.3%)	—	—
Age 61	8,078	5,636 (69.7%)	214 (2.6%)	2,228 (27.5%)	—	—
Age 62	7,951	4,058 (51.0%)	0	3,893 (49.0%)	3,741 (48.3%)	981
Age 63	7,762	1,657 (21.3%)	1,387 (17.8%)	4,718 (60.7%)	1,073 (13.8%)	1,155
Age 64	7,586	434 (5.7%)	2,413 (31.8%)	4,739 (62.5%)	815 (10.5%)	1,277
Age 65	7,420	175 (2.4%)	3,139 (42.3%)	4,106 (55.3%)	1,808 (23.3%)	1,390
Age 66	7,239	553 (7.6%)	4,179 (57.7%)	2,507 (34.6%)	306 (4.1%)	1,480
Model 3: ET with ARF Adjustments and Social Security Reform Uncertainty						
Age 60	8,234	5,768 (70.05%)	118 (1.43%)	2,348 (28.5%)	—	—
Age 61	8,078	5,636 (69.8%)	177 (2.2%)	2,265 (28.04%)	—	—
Age 62	7,951	3,377 (42.5%)	0 (0.00%)	4,574 (57.5%)	4,603 (59.08%)	997
Age 63	7,762	1,540 (19.9%)	2,000 (25.8%)	4,222 (54.4%)	813 (10.4%)	1,180
Age 64	7,586	394 (5.2%)	2,802 (36.93%)	4,390 (57.9%)	689 (8.8%)	1,275
Age 65	7,420	181 (2.4%)	3,511 (47.3%)	3,728 (50.2%)	1,659 (21.3%)	1,390
Age 66	7,239	594 (8.2%)	4,351 (60.1%)	2,294 (31.2%)	27 (0.4%)	1,514

Notes: ^aIn numbers, and as percentage of survivors. ^bNumber of First Claimers at that age, and as percentage of the total who ever claimed.

Table 4: Policy Experiment: NRA is now 67. 10,000 Simulations of the Dynamic Model

Ages	Survivors	Full-Time ^a	Part-Time ^a	No Work ^a	Claimers ^b	Benefits in \$
Model 1: Earnings Test as a Tax, and NRA=67						
Age 60	8,234	5,771 (70.1%)	117 (1.4%)	2,346 (28.5%)	—	—
Age 61	8,078	5,640 (69.8%)	166 (2.1%)	2,272 (28.1%)	—	—
Age 62	7,951	5,530 (69.5%)	0	2,423 (30.5%)	1,872 (24.8%)	989
Age 63	7,762	4,237 (54.6%)	194 (2.5%)	3,331 (42.9%)	446 (5.9%)	1,110
Age 64	7,586	1,744 (22.9%)	514 (6.7%)	5,328 (70.2%)	1,040 (13.78%)	1,212
Age 65	7,420	1,124 (15.1%)	0	6,298 (84.8%)	1,884 (24.9%)	1,234
Age 66	7,239	848 (11.7%)	0	6,393 (88.3%)	1,862 (24.6%)	1,392
Age 67	7,041	1,078 (15.3%)	0	5,968 (84.7%)	412 (5.46%)	1,457
Model 2: Earnings Test with ARF Adjustments, and NRA=67						
Age 60	8,234	5,773 (70.1%)	123 (1.5%)	2,338 (28.4%)	—	—
Age 61	8,078	5,636 (69.7%)	173 (2.1%)	2,269 (28.1%)	—	—
Age 62	7,951	4,150 (52.1%)	3 (0.03%)	3,798 (47.8%)	4,016 (51.96%)	886
Age 63	7,762	2,173 (27.99%)	1,409 (18.1%)	4,180 (53.85%)	996 (12.88%)	1,049
Age 64	7,586	839 (11.1%)	2,400 (31.6%)	4,347 (57.3%)	605 (7.8%)	1,182
Age 65	7,420	309 (4.16%)	3,024 (40.7%)	4,087 (55.1%)	738 (9.55%)	1,259
Age 66	7,239	158 (2.2%)	3,670 (50.69%)	3,411 (47.1%)	1,302 (16.84%)	1,397
Age 67	7,041	466 (6.6%)	4,314 (61.3%)	2,261 (32.1%)	71 (0.91%)	1,516
Model 3: ET with ARF Adjustments, Social Security Reform Uncertainty, and NRA=67						
Age 60	8,234	5,765 (70.01%)	118 (1.43%)	2,347 (28.5%)	—	—
Age 61	8,078	5,634 (69.75%)	177 (2.2%)	2,267 (28.06%)	—	—
Age 62	7,951	3,371 (42.4%)	0 (0.00%)	4,578 (57.57%)	5,084 (65.16%)	1,060
Age 63	7,762	2,016 (25.97%)	2,000 (25.8%)	3,597 (46.3%)	714 (9.15%)	1,073
Age 64	7,586	769 (10.1%)	2,802 (36.93%)	3,883 (51.1%)	512 (6.56%)	1,191
Age 65	7,420	316 (4.2%)	3,511 (47.3%)	3,552 (47.9%)	927 (11.9%)	1,281
Age 66	7,239	158 (2.2%)	4,392 (60.6%)	2,736 (37.8%)	561 (7.2%)	1,398
Age 67	7,041	499 (7.1%)	4,392 (60.6%)	2,217 (31.5%)	4 (0.05%)	1,523

Notes: ^aIn numbers, and as percentage of survivors. ^bNumber of First Claimers at that age, and as percentage of the total who ever claimed.

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.
- Benítez-Silva, H., and F. Heiland (2006): “Early Claiming of Social Security Benefits and Labor Supply Behavior of Older Americans,” forthcoming in *Applied Economics*.
- 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., M. Buchinsky, and J. Rust (2003): “Dynamic Structural Models of Retirement and Disability,” manuscript, SUNY-Stony Brook, UCLA, and University of Maryland.
- Benítez-Silva, H., M. Buchinsky, and J. Rust (2006): “Induced Entry Effects of a \$1 for \$2 Offset in SSDI Benefits,” manuscript, SUNY-Stony Brook, UCLA, and University of Maryland.
- Benítez-Silva, H., and N. Yin (2007): “Actuarial Fairness, Retirement Benefits, and Individual Heterogeneity,” manuscript, SUNY-Stony Brook.
- Blau, D. (2004): “Retirement and Consumption in a Life Cycle Model,” manuscript, University of North Carolina at Chapel Hill.
- 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.
- 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. Jouten (2002): “Delays in Claiming Social Security Benefits,” *Journal of Public Economics*, **84-3** 357–385
- De Nardi, Mariacristina, E. French, and J.B. Jones (2006): “Differential Mortality, Uncertain Medical Expenses, and the Saving of Elderly Singles,” manuscript.
- 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.
- French, E. (2005): “The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour,” *Review of Economic Studies*, **72** 395–427.
- 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.
- 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.
- 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.
- Judd, K.L. (1998): *Numerical Methods in Economics*, The MIT Press.
- 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.
- 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.
- Magnac, T. and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, **70-2** 801–816.
- Mastrobuoni, G. (2006): "Do better-informed workers make better retirement choices? A test based on the Social Security Statement," manuscript, Princeton University.
- Myers, R.J. (1993): *Social Security*. Fourth Edition. Pension Research Council and University of Pennsylvania Press.
- Phelan, C. (1999): "Anticipation effects of looming public-pension reforms: A comment." *Carnegie-Rochester Conference Series on Public Policy*, **50** 161–164.
- Phelan, C. (2006): "Public trust and government betrayal," forthcoming in *Journal of Economic Theory*.

- 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. (1994): "Structural Estimation of Markov Decision Processes," in R. Engle and D. McFadden (eds.) *Handbook of Econometrics*, Vol. 4, 3082–3139, North Holland.
- Rust, J. (1996): "Numerical Dynamic Programming in Economics," in *Handbook of Computational Economics*. H. Amman et al. (eds.). Amsterdam: Elsevier.
- Rust, J., and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, **65** 781–831.
- Sabelhauss, J., and J. Topoleski (2007): "Uncertain Policy for an Uncertain World: The Case of Social Security," *Journal of Policy Analysis and Management*, **26-3** 507–525.
- Song, J.G. (2004): "Evaluating the Initial Impact of Eliminating the Retirement Earnings Test," *Social Security Bulletin*, **65-1** 1–15.
- Song, J.G., and J. Manchester (2007): "New evidence on earnings and benefit claims following changes in the retirement earnings test in 2000," *Journal of Public Economics*, **91(3-4)** 669–700.
- SSA-S (various years): Annual Statistical Supplement to the Social Security Bulletin.
<http://www.ssa.gov/policy/docs/statcomps/supplement/>.
- SSA-H: Social Security Handbook. Online Version.
http://www.ssa.gov/OP_Home/handbook/ssa-hbk.htm.
- SSA-M: Program Operations Manual System (POMS). Public Version.
<https://s044a90.ssa.gov/apps10/poms.nsf/aboutpoms>.
- SSA (2007): *Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds*. Washington, D.C.
- Taber, C.R. (2000): "Semiparametric Identification and Heterogeneity in Discrete Choice Dynamic Programming Models," *Journal of Econometrics*, **96** 201–229.
- Votruba, M. E. (2003): "Social Security And Retirees' Decision to Work," manuscript, Case Western Reserve University.
- Vroman, W. (1985): "Some Economic Effects of the Retirement Test," in *Research in Labor Economics*, Vol. 7. JAI Press.

van der Klaauw, W., and K.I. Wolpin (2005): “Social Security, Pensions and the Savings and Retirement Behavior of Households,” manuscript. University of North Carolina-Chapel Hill.

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