Effects of social security policies on benefit claiming, retirement and saving☆

Alan L. Gustman [a],*, Thomas L. Steinmeier [b],[1]

[a] Department of Economics, Dartmouth College, Hanover, NH 03755, United States
[b] Department of Economics, Texas Tech University, Lubbock, TX 70409, United States

ABSTRACT

An enhanced version of a structural model jointly explains benefit claiming, wealth and retirement, including reversals from states of lesser to greater work. The model is estimated with Health and Retirement Study data. Alternative beliefs about the future of Social Security affect claiming behavior. Effects of three potential policies are also examined: increasing the early entitlement age, increasing the full retirement age, and eliminating the payroll tax for seniors. Predicted responses to increasing the full entitlement age are sensitive to beliefs.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Social Security continues to be under substantial financial pressure (Social Security Administration Trustees Report, 2014). A number of policy changes have been suggested to encourage a population that is becoming increasingly long lived to work longer, thereby improving the financial status of the Social Security.[2] Increasing the early entitlement age for benefits, increasing the full retirement age, and abolishing the payroll tax on earnings for those over the full retirement age are among the policy changes that are frequently mentioned. These policy changes will affect the timing of retirement (withdrawal from the labor market), the resources available to seniors when they leave the labor market, and the timing of benefit claiming.

To analyze the effects of these potential changes in policy, this paper specifies and estimates a structural model that is designed to simultaneously explain retirement outcomes, saving (wealth accumulation) and benefit claiming. This model updates our previous work on a number of dimensions, most importantly by including reversals from states of lesser to greater work. The model also updates our baseline model to include stochastic returns on assets. These innovations allow us to explain the full range of retirement dynamics, heterogeneity in wealth accumulation, and benefit claiming, an outcome that researchers have had particular problems in explaining.

We then introduce different beliefs about the future of Social Security and trace the effects of these beliefs on claiming, retirement and saving. Three sets of beliefs are incorporated in alternative simulations and compared to outcomes under a baseline scenario. In the base case, covered workers believe that they will receive benefits as called for under current law. In alternative scenarios individuals believe that Social Security benefits will be reduced in the future; they believe that the return on assets will be higher in the future than in the present; or they believe that survivor benefits are less valuable than own benefits. Although we do not incorporate each of the alternative beliefs in the model estimation, we do incorporate the alternative beliefs in simulations that allow us to place bounds on their likely influences on key outcomes.

Our estimates are based on data from the Health and Retirement Study. We utilize detailed employer pension plan descriptions, which

allows us to include sharp discontinuities in the budget set that vary according to the relevant plan description and each individual's employment history with the firm. We also utilize matched earnings histories obtained from the Social Security Administration.

Section 2 provides an overview of the retirement literature, including our own related contributions, and discusses some outstanding issues in the literature. A structural model of retirement, saving and claiming is developed in Section 3. The model is estimated and its properties are examined through simulation in Section 4. Section 5 compares this model with models we have estimated in our previous studies. It provides measures of goodness of fit in explaining exits from full-time work and exits from the labor market. Section 6 explores modifications of the claiming model that would explain the apparently excessive claiming observed at younger ages. There we focus on differences in beliefs about the future of own benefits and understanding of the value of spouse benefits. The effects of the three key policies, increasing the early entitlement, increasing the full retirement age, and abolishing the payroll tax for those over the full retirement age, are examined in Section 7. Section 8 compares predictions of the effects of increasing the early entitlement age across five versions of our retirement model. Section 9 concludes.

2. Overview of the literature relating Social Security to retirement, benefit claiming and saving

Until the last couple of decades, the most prominent feature of the retirement hazard was the spike in retirements at the Social Security full retirement age. This spike was visible in both national and international data (Gruber and Wise, 1999). It was often attributed to unfair actuarial adjustments for work after full retirement age. When a person delayed retiring, Social Security and pension benefits were not paid. Benefits to be paid in the future were not increased at an actuarially fair rate to compensate for benefit payments lost to any earnings test while continuing at work. In the extreme, when benefits were not increased at all or by a very small amount, as when the Social Security delayed retirement credit was 1%, a person might find the net wage for work after qualifying for full retirement benefits to have fallen to the value of the nominal wage paid before retirement, less the value of the foregone pension or Social Security benefit. Moreover, mandatory retirement was legal and common. Both factors accounted for much of the spike in retirement at full retirement age.

That all has changed. Social Security legislation increased the delayed retirement credit and abolished the earnings test after the full retirement age. Other legislation abolished mandatory retirement, and related court decisions mandated fairer actuarial adjustments in pensions for older persons. These changes, along with other trends, substantially reduced the spike in retirements around the full retirement age (Anderson et al., 1999; Gustman and Steinmeier, 2009).

Today the spike in retirements at age 62 is the more prominent feature of the retirement hazard. That spike is almost certainly related to the Social Security early entitlement age. As would be expected, the spike in exits from the labor force at age 62 is accompanied by a spike in benefit claiming at that age.[3]

There is a puzzle, however. For most individuals when benefit receipt is postponed beyond age 62, future Social Security benefits are increased at better than an actuarially fair rate.[4] A task of retirement and claiming models is to explain why many who have already retired do not defer claiming beyond age 62, and why many others do not defer retirement and claiming for a number of years after they reach age 62.

Heterogeneity in time preference rates has helped to explain the spike in early retirements at age 62 (Gustman and Steinmeier, 2005a). A question we explore here is whether heterogeneity in time preference can also help to explain claiming behavior. That is, where a substantial subgroup of the population has high discount rates, along with an inability to borrow from future income, do they account both for the spike in retirement at age 62, and for what in some models without heterogeneity in time preference appears to be an excess of benefit claiming at age 62? Those with a high time preference rate and little or no pension income would also be expected to have low levels of other assets (perhaps excluding their home). Although they might like to retire before age 62, they do not have sufficient assets to allow them to do so. At 62, when Social Security benefits become available, they retire and claim their benefits en masse. Others, with intermediate time preferences, may have accumulated some assets (Coile et al., 2002), but for them the prospect of increased future benefits is insufficient to compensate for delayed current benefits. These individuals will also want to collect benefits as soon as possible. Further, if they are forced to delay benefits because of an earnings test, the earnings test will serve to reduce their perceived compensation, making them more likely to retire at age 62 when the earnings test becomes effective. Thus, spikes in retirement and claiming at 62 are not necessarily limited to those who are asset constrained.

Heterogeneity in time preference rates also helps to explain the wide variation in wealth, even for those with similar levels of lifetime earnings. Venti and Wise (1999, 2001) find an extraordinarily wide distribution of wealth holdings. Of greatest interest, the distribution of wealth holding is very wide even among individuals who fall within the same decile of lifetime income. (See also Gustman and Steinmeier, 1999.) Evidence from the HRS also shows that there are a distressing number of age 50 + households with essentially no retirement savings outside of Social Security. Among those with no savings outside of Social Security, many have had substantial earnings in the past.

Scholz et al. (2006) claim that the wealth distribution at retirement can be pretty well explained without time preference heterogeneity. In their model, almost no one has wealth that is significantly lower than their optimal amount (see their Fig. 2). Yet Venti and Wise find a large number of individuals, even at fairly high income levels, who have very little wealth, generating very wide wealth distributions for those falling within the same lifetime income decile. Moreover, in their Fig. 3, Scholz et al. measure the discrepancy between actual and optimal net worth. That figure indicates that for a given income decile, there is a substantial variation in the difference between actual net worth and their calculated optimal net worth. This suggests that their model with uniform time preference does not do as good a job in generating the variation of observed assets within income deciles as might appear.[5]

Other possible explanations of low net worth for high income individuals are also suspect. One common explanation is stochastic wages. If current wages are higher than expected, then the savings which were done on the expectation of lower wages may appear small relative to current wages. However, this would not explain assets so low as to be almost negligible, and in any case unless the unexpected wage changes were very recent, savings and asset levels would adjust to the higher wage levels over time. A similar argument is that unexpected negative asset returns can cause low assets relative to savings (think Enron). But once again, unless the negative asset returns were very recent,

---

[3] Card et al. (2014) document a modest decline in the age 62 spike in claiming over recent years.

[4] It has been known for some time that at a 3% real interest rate, postponing benefit receipt, especially in the first few years after age 62, increases future benefits for couples at better than an actuarially fair rate (Gordon and Blinder, 1980; Feldstein and Samwick, 1992). More recently, Shoven and Slavov (2012a,b) show that at the very low market interest rates following the Great Recession, delayed claiming is an even better deal.

[5] On page 626, Scholz et al. claim that their model explains 86% of the observed variation in net worth. This result may be related to the concentration of observations in the lower left corner of the scatter plot in their Fig. 2. These observations correspond to lower-income individuals with low values of both observed and calculated optimal net worth. For low income individuals, these two values must necessarily be fairly close.