Limited Computational Ability and Social Security

Frank N. Caliendo T. Scott Findley*

September 11, 2009

Abstract

We compute the optimal social security tax rate in a calibrated general equilibrium model with agents who lack the computational ability to solve dynamic optimization problems. Instead, they follow the simple rule of thumb of consuming and saving a fixed fraction of disposable income. This departs from a literature that computes the optimal provision of social security when agents are mathematically sophisticated but suffer from some behavioral bias (hyperbolic preferences, temptation preferences, short planning horizons, etc.). The calibrated model replicates the aggregate capital-output ratio, the mean saving rate in the US, and the fraction of individuals who experience a drop in consumption spending at the date of retirement. Thus, the calibrated model matches aggregate and micro-level indicators of preparedness for retirement. Our quantitative results are generally supportive of a social security program as large as the one in the US.

Keywords: rule of thumb, consumption, saving, optimal social security, general equilibrium calibration.

JEL Classifications: H55, D91, C61

*We thank the participants of the 2009 Summer Workshop of the *Quantitative Society for Pensions and Saving* for comments. Both authors are in the Jon M. Huntsman School of Business at Utah State University. Send correspondence to Findley at: tscott.findley@usu.edu.

1 Introduction

Pay-as-you-go social security programs are common in industrialized countries and they (arguably) serve a variety of functions. They may help people save for retirement; they transfer wealth from the elderly wealthy to the elderly poor; they protect against the risk of outliving one's assets; and, they provide of variety of other insurance roles relating to disability and survivorship. This paper follows a literature that focuses on the first function.

Kotlikoff, Spivak, and Summers (1982) say the "essential premise underlying the social security program...is that left to their own devices, large numbers of people would fail to save adequately and find themselves destitute in their old age." At first glance, it may be tempting to conclude that social security can easily be justified on this rationale since about one-third of retirees have no income source other than social security, and social security is the largest income source for two-thirds of retirees. However, it is well known that social security with a below market internal rate of return cannot improve the lifetime welfare of a standard life-cycle permanent-income (LCPI) consumer, no matter how impatient the consumer may be (e.g., see Findley and Caliendo 2008).

Beginning with Feldstein (1985), economists have therefore asked whether standard life-cycle models augmented to include some type of short-sightedness or myopia can be used to rationalize a pay-as-you-go program. The trend has been to keep the standard assumption that people are *optimizers*, though they may solve the "wrong" optimization problem or suffer from some type of self control problem. For example, İmrohoroğlu, İmrohoroğlu, and Joines (2003) consider how sophisticated quasi-hyperbolic consumers (Laibson 1997) fare in a social security program. Kumru and Thanopoulos (2008) and Bucciol (2008) focus on the same policy issue, but instead assume that consumers have self control problems like in Gul-Pesendorfer (2001). Findley and Caliendo (2009) also study the same policy issue, but for consumers who have short planning horizons. The results of these studies are highly mixed and usually depend critically on the particular values of the preference parameters used in the calibration procedure, with policy implications ranging from complete abolition of the social security program to a doubling of its size.[1]

[1] We provide surveys of this literature in a few of our companion papers. See Findley and Caliendo (2008)

In each of these papers (İmrohoroğlu, İmrohoroğlu, and Joines 2003, Kumru and Thanopoulos 2008, Bucciol 2008, Findley and Caliendo 2009) the math problem solved by the representative consumer is at least as difficult as the standard LCPI problem, and in some cases the math problem is an order of magnitude more sophisticated. Of course, this is a common critique of behavioral economics. In trying to make the consumer somewhat boundedly rational and therefore more like a real person, the complexity of the math problem solved by the consumer in the model blows up. With this is mind, we want to calculate the optimal social security tax rate in a model where the consumer lacks the computational ability to solve complex math problems, and instead follows a simple rule of thumb.

Studying the optimal provision of social security for rule-of-thumb consumers is not an academic exercise. Casual observation and other evidence suggest that rule-of-thumb behavior could be widespread, and is therefore relevant to this important policy discussion. In the standard LCPI model with perfect foresight, the agent calculates a detailed plan for retirement consumption immediately upon entering the workforce. Survey evidence generally does not seem to support this assumption and is arguably more in line with rule-of-thumb behavior.[2] According to Yakoboski and Dickemper (1997), 64 percent of households in the 1997 Retirement Confidence Survey have never tried to calculate how much they need to save for retirement. Similarly, Ameriks, Caplin, and Leahy (2003) ask a survey of TIAA-CREF participants the following question: "Have you personally gathered together your household's financial information, reviewed it in detail, and formulated a specific financial plan for your household's long term future?" Even in this survey of highly educated individuals, who by definition save for retirement, 27 percent answer "no" to this question. In addition, one-third of the respondents from the Health and Retirement Study (people aged 51-61) report that

and Findley and Caliendo (2009).

[2]Of course, some people may indeed solve optimization problems (or at least behave as if they do). But behavioral economists (e.g., Thaler 1994 and many others) have long contended that the life-cycle problem, or virtually any variant of it, is hard to solve and most households have limited computational powers. And although repetition helps people learn to solve difficult problems, there may be few opportunities for learning when it comes to saving for retirement. It is not a repeat game. You only get one shot at it. Although, a counter argument is that people learn by watching others (Lusardi 2003) and by listening to professionals (Thaler 1994).

they have "hardly thought about retirement." Of course, there are multiple ways to interpret this self-reported data, but it clearly seems inconsistent with the hypothesis that people are dynamic optimizers who make explicit consumption plans that stretch all the way out to the date of death. The evidence seems more consistent with rule-of-thumb behavior which does not necessarily involve explicit retirement planning.

The closest concept to a lack of computational power in the related social security literature is hand-to-mouth behavior (Feldstein 1985, Docquier 2002, Cremer, De Donder, Maldonado, and Pestieau 2008, Caliendo and Gahramanov 2009). However, these models assume a discrete distribution: either people are perfect planners who solve an optimization problem or else they just consume all of their income. The standard rule-of-thumb approach in which people save a constant fraction of disposable income is a natural middle point.

We build a continuous-time overlapping-generations model with consumers who save a constant fraction of their disposable wage income during the working years and then withdraw a constant annuity stream during retirement. The government operates a pay-as-you-go social security program which is "Bismarckian" in nature (Cremer, De Donder, Maldonado, and Pestieau 2008). Individuals receive social security benefits that are proportional to their productivity during the working period. Thus, we abstract from the redistributive component of a social security program to isolate the role of social security in helping people save for retirement. We consider general equilibria in which individuals differ by age, productivity, and saving rates. Following convention in this literature (İmrohoroğlu, İmrohoroğlu, and Joines 2003, Kumru and Thanopoulos 2008, Bucciol 2008, Findley and Caliendo 2009, Caliendo and Gahramanov 2009), we calibrate the model to match the aggregate capital-output ratio. Our calibrated model also approximately replicates the mean saving rate in the US and the fraction of individuals who experience a discrete drop in consumption spending at the date of retirement. Thus, our calibrated model matches aggregate and micro-level indicators of preparedness for retirement and therefore appears to be a reasonable instrument for studying the question at hand.

We use the calibrated model to compute the optimal social security tax rate for a variety of assumptions about the preferences of a benevolent social planner and for a variety of assumptions about the distribution of saving rates in the economy. If we adopt the Ramsey

(1928) ideology in which the social planner does not discount future utility, then the optimal size of the social security program in our model is very close to the current size of the program in the US.[3] This result is robust to moderate changes in the distribution of saving rates.

Our results contrast sharply with İmrohoroğlu, İmrohoroğlu, and Joines (2003), which is the pathbreaking paper for studying the optimal provision of social security in general-equilibrium overlapping-generations models with bounded rationality. Their paper has become the standard upon which the latter general equilibrium papers base their comparison (Kumru and Thanopoulos 2008, Bucciol 2008, Findley and Caliendo 2009, Caliendo and Gahramanov 2009). İmrohoroğlu, İmrohoroğlu, and Joines prove analytically that hyperbolic preferences alone cannot rationalize the existence of a social security program with a below market internal rate of return. In fact, the only way to get a positive optimal social security tax rate in their model is if the general equilibrium distortions to factor prices happen to be just right. Therefore, social security is not viewed as a useful commitment device. It is only a welfare improving program under a fragile set of circumstances. In contrast, social security plays a significant welfare-improving role in our study for a wide range of assumption about the parameter values, whether in partial equilibrium or general equilibrium.

2 Life-cycle, general equilibrium model with rule-of-thumb consumers

Time is continuous and is indexed by t. We consider an individual who is born at $t = \tau$, retires at $t = \tau + T$, and passes away at $t = \tau + \bar{T}$. The consumer saves a fixed fraction s of disposable wage income during the working years and annuitizes the savings balance upon retirement. Hence the consumer does not need to know how to solve a complex math problem. He simply saves a fixed fraction of his wages and then asks his pension provider to give him a smooth payout over the retirement period. The economy-wide wage per unit of labor at time t is $w(t) = w(\tau)e^{x(t-\tau)}$, where x is the rate of real wage growth. The individual's

[3] Portney and Weyant (1999) compile a series of papers relating to the optimal social discount rate in long-term public projects. Although not all authors share the same view, a number of them argue that zero discounting is a fair approach.

productivity is φ, and so he earns $\varphi w(t)$ at time t. The productivity parameter φ varies across individuals but is constant across the worklife of a given individual. Social security taxes are the only taxes on wage income, and they occur at rate θ. Following convention in this literature, the worker bears the full burden of the tax.

The individual's savings account, $k(t, \tau|s, \varphi)$, evolves according to the following differential equations

$$\frac{dk(t, \tau|s, \varphi)}{dt} = rk(t, \tau|s, \varphi) + s(1-\theta)\varphi w(t), \quad \text{for } t \in [\tau, \tau + T], \tag{1}$$

$$\frac{dk(t, \tau|s, \varphi)}{dt} = rk(t, \tau|s, \varphi) - a(\tau|s, \varphi), \quad \text{for } t \in [\tau + T, \tau + \bar{T}], \tag{2}$$

where r is the real rate of return on private saving, and $a(\tau|s, \varphi)$ is the constant annuity withdrawal that drives the savings account to zero by the end of the life cycle. Using the boundary conditions, $k(\tau, \tau|s, \varphi) = 0$ and $k(\tau + \bar{T}, \tau|s, \varphi) = 0$, we have

$$k(t, \tau|s, \varphi) = \int_{\tau}^{t} s(1-\theta)\varphi w(z)e^{r(t-z)}dz, \quad \text{for } t \in [\tau, \tau + T], \tag{3}$$

$$k(t, \tau|s, \varphi) = \int_{t}^{\tau+\bar{T}} a(\tau|s, \varphi)e^{r(t-z)}dz, \quad \text{for } t \in [\tau + T, \tau + \bar{T}]. \tag{4}$$

Evaluating (3) and (4) at $t = \tau + T$ and solving for $a(\tau|s, \varphi)$ gives

$$a(\tau|s, \varphi) = \frac{\displaystyle\int_{\tau}^{\tau+T} s(1-\theta)\varphi w(t)e^{r(\tau+T-t)}dt}{\displaystyle\int_{\tau+T}^{\tau+\bar{T}} e^{r(\tau+T-t)}dt}, \quad \text{for } t \in [\tau + T, \tau + \bar{T}]. \tag{5}$$

The numerator of (5) is the balance in the savings account at the date of retirement, and the denominator is the value at the date of retirement of a one-dollar annuity over the retirement period. Note that the individual does not need to compute (5) by himself, since virtually any defined contribution pension provider would readily convert his account balance at retirement to an annuity. It is convenient to integrate (3)-(5)

$$k^1(t, \tau|s, \varphi) = j_1 s\varphi w(\tau)e^{x(t-\tau)} - j_1 s\varphi w(\tau)e^{r(t-\tau)}, \tag{6}$$

$$k^2(t, \tau|s, \varphi) = j_2 s\varphi w(\tau) - j_3 s\varphi w(\tau)e^{r(t-\tau)}, \tag{7}$$

where $k^1(t,\tau|s,\varphi)$ represents the work-life savings account path and $k^2(t,\tau|s,\varphi)$ is the path during retirement, and

$$j_1 \equiv \frac{1-\theta}{x-r}, \tag{8}$$

$$j_2 \equiv \frac{(1-\theta)\left[e^{(x-r)T}-1\right]}{(x-r)\left[e^{-rT}-e^{-r\bar{T}}\right]}, \tag{9}$$

$$j_3 \equiv j_2 e^{-r\bar{T}}. \tag{10}$$

The agent's life-cycle consumption path is piecewise continuous

$$c(t,\tau|s,\varphi) = (1-s)(1-\theta)\varphi w(t), \quad \text{for } t \in [\tau, \tau+T], \tag{11}$$

$$c(t,\tau|s,\varphi) = a(\tau|s,\varphi) + b(t), \quad \text{for } t \in [\tau+T, \tau+\bar{T}], \tag{12}$$

where $b(t)$ is the social security benefit received during retirement.

At each instant a new cohort is born and the oldest cohort dies, and each cohort contains a mass of infinitely divisible people. The number of individuals born at time τ is $N(\tau) = N(t)e^{n(\tau-t)}$. Let worker productivity φ be continuous and uniformly distributed with support $[\varphi_0, 1]$ and density $f(\varphi) = (1-\varphi_0)^{-1}$ within every cohort, where $0 \le \varphi_0 \le 1$. This normalization implies that the most productive worker is a "full worker", while all other workers add just some fraction φ of the effort rendered by the most productive worker in the production process. The number of *individuals* who are of working age at time t is $\int_{t-T}^{t} N(t)e^{n(\tau-t)}d\tau$, but the total supply of *labor* at time t is

$$L(t) = \int_{t-T}^{t}\int_{\varphi_0}^{1} \varphi f(\varphi)N(t)e^{n(\tau-t)}d\varphi\, d\tau = \int_{t-T}^{t} \frac{1+\varphi_0}{2} N(t)e^{n(\tau-t)}d\tau. \tag{13}$$

The last expression in (13) shows that the labor supply is equal to the total number of people who are of working age multiplied by average productivity. Note that n is both the rate of growth in cohort size as well as the rate of growth in the labor supply so that $L(t) = L(0)e^{nt}$.

Aggregate capital, $K(t)$, is the sum of micro-level demand for capital across and within each cohort, where each cohort contains heterogeneity with respect to the saving rate and productivity. All of the heterogeneity is *within* a given cohort (not across different cohorts), meaning that all cohorts are identical after controlling for technological growth and population growth. We assume that the saving rate is continuous and uniformly distributed within every cohort with support $[s_0, s_1]$ and density $g(s) = (s_1 - s_0)^{-1}$. Note that the life-cycle capital profile is not differentiable at the instant of retirement (there is a kink). Therefore,

$$\begin{aligned} K(t) &= \int_{t-T}^{t} \int_{s_0}^{s_1} \int_{\varphi_0}^{1} f(\varphi) g(s) N(t) e^{n(\tau - t)} k^1(t, \tau | s, \varphi) \, d\varphi \, ds \, d\tau \\ &+ \int_{t-\bar{T}}^{t-T} \int_{s_0}^{s_1} \int_{\varphi_0}^{1} f(\varphi) g(s) N(t) e^{n(\tau - t)} k^2(t, \tau | s, \varphi) \, d\varphi \, ds \, d\tau. \end{aligned} \tag{14}$$

We have assumed s and φ are independently distributed, or that for every possible saving rate there exists a continuum of productivities that are spread evenly between φ_0 and 1, and likewise for every φ there exists a continuum of saving rates spread evenly between s_0 and s_1. With (6) and (7) in mind, we can multiply each integrand in (14) by $\frac{s\varphi}{s\varphi}$ and rewrite

$$\begin{aligned} K(t) &= \int_{t-T}^{t} \frac{k^1(t, \tau | s, \varphi)}{s\varphi} N(t) e^{n(\tau - t)} \int_{s_0}^{s_1} \int_{\varphi_0}^{1} f(\varphi) g(s) s\varphi \, d\varphi \, ds \, d\tau \\ &+ \int_{t-\bar{T}}^{t-T} \frac{k^2(t, \tau | s, \varphi)}{s\varphi} N(t) e^{n(\tau - t)} \int_{s_0}^{s_1} \int_{\varphi_0}^{1} f(\varphi) g(s) s\varphi \, d\varphi \, ds \, d\tau \\ &= \int_{t-T}^{t} \frac{k^1(t, \tau | s, \varphi)}{s\varphi} N(t) e^{n(\tau - t)} \left(\frac{s_0 + s_1}{2} \right) \left(\frac{1 + \varphi_0}{2} \right) \, d\tau \\ &+ \int_{t-\bar{T}}^{t-T} \frac{k^2(t, \tau | s, \varphi)}{s\varphi} N(t) e^{n(\tau - t)} \left(\frac{s_0 + s_1}{2} \right) \left(\frac{1 + \varphi_0}{2} \right) d\tau \\ &= \int_{t-T}^{t} k^1 \left(t, \tau | \tfrac{s_0 + s_1}{2}, \tfrac{1 + \varphi_0}{2} \right) N(t) e^{n(\tau - t)} \, d\tau \\ &+ \int_{t-\bar{T}}^{t-T} k^2 \left(t, \tau | \tfrac{s_0 + s_1}{2}, \tfrac{1 + \varphi_0}{2} \right) N(t) e^{n(\tau - t)} \, d\tau, \end{aligned}$$

which clearly shows that the aggregate demand for capital for cohort τ (the integrand) is just equal to the demand for capital of the average agent from cohort τ multiplied by the size of cohort τ. Integrating gives

$$K(t) = N(t) w(t) \left(\frac{s_0 + s_1}{2} \right) \left(\frac{1 + \varphi_0}{2} \right) \Gamma, \tag{15}$$

where Γ is independent of birth and calendar dates

$$\begin{aligned} \Gamma \equiv \; & \frac{j_1}{n} \left[1 - e^{-nT} \right] + \frac{j_1}{n + x - r} \left[e^{(r - x - n)T} - 1 \right] \\ & + \frac{j_2}{n + x} \left[e^{-(n + x)T} - e^{-(n + x)\bar{T}} \right] \\ & + \frac{j_3}{n + x - r} \left[e^{(r - x - n)\bar{T}} - e^{(r - x - n)T} \right]. \end{aligned} \tag{16}$$

Output is Cobb-Douglas, $Y(t) = K(t)^\alpha [A(t)L(t)]^{1-\alpha}$, where $A(t) = A(0)e^{xt}$ is the stock of labor augmenting technology and α is capital's share in income. Factors are priced competitively

$$\begin{aligned} r &= Y_K(t) - \delta, & (17) \\ w(t) &= Y_L(t), & (18) \end{aligned}$$

where δ is the rate of capital depreciation. Note that $w(t)$ is the wage per unit of labor, so only the most productive worker with $\varphi = 1$ earns $w(t)$, everyone else earns $\varphi w(t) \in [\varphi_0 w(t), w(t))$. For constant returns technology, Euler's theorem ensures that total factor payments plus depreciation must equal total output

$$(r + \delta)K(t) + w(t)L(t) = Y(t). \tag{19}$$

Let us consider two separate cases of the definition of social security benefits. These two cases help us to disentangle the welfare effect of social security that is due to undersaving in isolation, versus the welfare effect that is due to both undersaving and income redistribution in tandem.

Case 1. Perfect Redistribution. In this case, social security provides protection against undersaving and it also serves to spread around wealth to reduce inequality among the elderly. This is the so-called Beveridgean pension system (Cremer, De Donder, Maldonado, and Pestieau 2008). All retirees alive at time t earn a social security benefit $b(t)$, regardless of their earnings histories, where a balanced budget requires that total taxes paid at time t equals total benefits received at time t

$$b(t) = \frac{\theta w(t)L(t)}{\int_{t-\bar{T}}^{t-T} N(t)e^{n(\tau-t)}d\tau}. \tag{20}$$

Although all retirees receive the same social security benefits, heterogeneity in saving rates and earnings can lead to substantial heterogeneity in the private annuity, $a(\tau|s,\varphi)$. This type of system would insure against undersaving and also reduce income inequality among retirees at the expense of increasing inequality in the internal rate of return experienced over the life cycle.

Case 2. Perfect benefit-earning correlation (i.e., no redistribution). In this case, the only function of social security is to provide protection against undersaving. This is the so-called

Bismarkian pension system (Cremer, De Donder, Maldonado, and Pestieau 2008). Benefits now depend on individual productivity as well as aggregate technology, $b(t|\varphi)$, ranging from a low value of $b(t|\varphi_0)$ to a high value of $b(t|1)$. To shut off the redistributive aspect of the social security program, it must be the case that the ratio of highest to lowest benefits equals the ratio of highest to lowest productivities,

$$\frac{b(t|1)}{b(t|\varphi_0)} = \varphi_0^{-1}. \tag{21}$$

In general the level of benefits, $b(t|\varphi)$, must equal $\varphi b(t|1)$, and therefore must lie on the straight line (rather than a curve, which would involve redistribution) connecting the ordered pairs $(\varphi_0, b(t|\varphi_0))$ and $(1, b(t|1))$. Expressed in point-slope form,

$$b(t|\varphi) = b(t|\varphi_0) + \frac{b(t|1) - b(t|\varphi_0)}{1 - \varphi_0} \times (\varphi - \varphi_0). \tag{22}$$

Combining (21) and (22) gives

$$b(t|\varphi) = b(t|\varphi_0)\varphi_0^{-1}\varphi. \tag{23}$$

Invoking the balance budget requirement gives

$$\theta w(t)L(t) = \int_{t-\bar{T}}^{t-T} \int_{\varphi_0}^{1} b(t|\varphi_0)\varphi_0^{-1}\varphi f(\varphi)N(t)e^{n(\tau-t)}d\varphi \; d\tau, \tag{24}$$

which clearly defines the benefits of the least productive member of society

$$b(t|\varphi_0) = \frac{\theta w(t)L(t)}{\displaystyle\int_{t-\bar{T}}^{t-T} \int_{\varphi_0}^{1} \varphi_0^{-1}\varphi f(\varphi)N(t)e^{n(\tau-t)}d\varphi \; d\tau}, \tag{25}$$

and hence, we can now write benefits in closed form

$$b(t|\varphi) = \frac{\theta w(t)L(t)\varphi_0^{-1}\varphi}{\displaystyle\int_{t-\bar{T}}^{t-T} \int_{\varphi_0}^{1} \varphi_0^{-1}\varphi f(\varphi)N(t)e^{n(\tau-t)}d\varphi \; d\tau}, \text{ for all } \varphi \in [\varphi_0, 1]. \tag{26}$$

Thus, once φ_0 has been specified and $w(t)$ has been computed in general equilibrium, the line segment $b(t|\varphi)_{\varphi\in[\varphi_0,1]}$ is exactly identified. Unlike Case 1, where there is an inverse relationship between individual productivity and the internal rate of return from social security, all agents experience the same internal rate of return. Hereafter, we focus on Case 2 in order to isolate the role of social security in helping people save for retirement.[4]

We limit our attention to **stable competitive equilibria** that simultaneously satisfy the following conditions: (i) all consumers choose consumption according to the simple rule of thumb, (11) and (12), taking factor prices as given; (ii) factors are priced competitively according to (17) and (18), where the interest rate is constant over time and the wage rate grows at a constant rate x; (iii) aggregate capital is the sum of household-level demand for capital according to (14); (iv) the population grows at a constant rate n; and, (v) the budget in the social security program balances at each instant and the level of benefits paid to each retiree at time t satisfies (26). Finding the equilibrium is a matter of finding the value of $K(t)$ that solves (14), which is an implicit function of $K(t)$ since factor prices depend on $K(t)$.[5] We resort to numerical approximation.

A key assumption in our analysis is that the individual deviates from the standard LCPI rule because he can't solve complex math problems, not because he is guilty of some sort of behavioral defect such as hyperbolic preferences, temptation, short planning horizons, or other types of self control problems. Hence, his true lifetime utility function is the same as in a standard LCPI problem

$$\begin{aligned} V(\tau|s,\varphi) \equiv & \int_{\tau}^{\tau+T} e^{-\rho(t-\tau)} \frac{\left[(1-s)(1-\theta)\varphi w(t)\right]^{1-\sigma}}{1-\sigma} dt \\ & + \int_{\tau+T}^{\tau+\bar{T}} e^{-\rho(t-\tau)} \frac{\left[a(\tau|s,\varphi)+b(t|\varphi)\right]^{1-\sigma}}{1-\sigma} dt, \end{aligned} \tag{27}$$

where ρ is the discount rate, σ is the inverse elasticity of intertemporal substitution, given

[4] See Cremer, De Donder, Maldonado, and Pestieau (2008) and Caliendo and Gahramanov (2009) for a detailed analysis of optimal social security when the redistribution mechanism is active.

[5] While the type of social security benefit rule (be it Case 1 or Case 2) may have important implications for lifetime welfare, it does not affect the aggregate equilibrium of the economy. The agents are not forward looking and so their demand for capital during the working phase does not depend on the level of benefits to be received later in retirement, and the demand for capital during retirement is also invariant to benefits because the agent is just following a simple rule of running down his existing savings. However, the size of the social security tax rate certainly affects aggregate quantities because demand for capital in the working phase is inversely related to θ.

the benefit rule for Case 2 with no redistribution. Note that

$$V(\tau|s,\varphi) = \varphi^{1-\sigma}V(\tau|s,1). \tag{28}$$

We measure social welfare for a representative cohort as[6]

$$S(\tau) \equiv \int_{s_0}^{s_1}\int_{\varphi_0}^{1} f(\varphi)g(s)V(\tau|s,\varphi)d\varphi\ ds = \Omega\int_{s_0}^{s_1} g(s)V(\tau|s,1)\ ds, \tag{29}$$

where $\Omega \equiv \int_{\varphi_0}^{1} f(\varphi)\varphi^{1-\sigma}d\varphi$. By definition, the optimal social security tax maximizes $S(\tau)$.

3 Some theoretical results

Lemma 1. *The capital-output ratio, the output-labor ratio, the interest rate, and the wage rate are all invariant to* φ_0.

Proof. *For a given interest rate and wage rate, the capital-output ratio*

$$\begin{aligned}\frac{K(t)}{Y(t)} &= \frac{K(t)}{K(t)^{\alpha}\left[A(t)\int_{t-T}^{t}\frac{1+\varphi_0}{2}N(t)e^{n(\tau-t)}d\tau\right]^{1-\alpha}} \\ &= \frac{\left[N(t)w(t)\left(\frac{s_0+s_1}{2}\right)\Gamma(r)\right]^{1-\alpha}}{\left[A(t)\int_{t-T}^{t}N(t)e^{n(\tau-t)}d\tau\right]^{1-\alpha}},\end{aligned}$$

and the output-labor ratio

$$\frac{Y(t)}{L(t)} = \left[\frac{K(t)}{L(t)}\right]^{\alpha} A(t)^{1-\alpha} = \left[\frac{N(t)w(t)\left(\frac{s_0+s_1}{2}\right)\Gamma(r)}{\int_{t-T}^{t}N(t)e^{n(\tau-t)}d\tau}\right]^{\alpha} A(t)^{1-\alpha},$$

are invariant to φ_0. *On the other hand for a given capital-output ratio and a given output-labor ratio, the interest and wage rates*

$$\begin{aligned} r &= Y_K(t) - \delta = \alpha Y(t)/K(t) - \delta, \\ w(t) &= Y_L(t) = (1-\alpha)Y(t)/L(t),\end{aligned}$$

[6] Note that $S(\tau)$ equals aggregate utility for cohort τ normalized by the size of the cohort, $N(\tau)$. Hence $S(\tau)$ is average individual welfare.

are also invariant to φ_0. In equilibrium, the capital-output ratio and the output-labor ratio are determined by factor prices, while factor prices are simultaneously determined by the capital-output ratio and the output-labor ratio. Hence, even though a change in φ_0 will cause a change in the equilibrium capital stock, it will not affect the capital-output ratio, the output-labor ratio, or factor prices. Q.E.D.

Lemma 2. *The private annuity, $a(\tau|s,1)$ and the social security benefit from Case 2, $b(t|1)$, are both invariant to φ_0.*

Proof. *The annuity $a(\tau|s,1)$ is not a function of φ_0 because Lemma 1 tells us that factor prices are not affected by φ_0. For the same reason, the social security benefit from (26)*

$$b(t|1) = \frac{\theta w(t) L(t)}{\frac{N(t)}{n}\frac{1+\varphi_0}{2}\left(e^{-nT} - e^{-n\bar{T}}\right)} = \frac{\theta w(t) \int_{t-T}^{t} N(t) e^{n(\tau - t)} d\tau}{\frac{N(t)}{n}\left(e^{-nT} - e^{-n\bar{T}}\right)}$$

is also not a function of φ_0. Q.E.D.

Proposition 1. *Under the benefit rule of Case 2, the optimal social security tax rate is invariant to φ_0.*

Proof. $\theta^* \equiv \arg\max \left[\Omega \int_{s_0}^{s_1} g(s) V(\tau|s,1)\ ds\right] = \arg\max \left[\int_{s_0}^{s_1} V(\tau|s,1)\ ds\right]$*, but using Lemmas 1 and 2, it is clear that $V(\tau|s,1)$ is not a function of φ_0. Q.E.D.*

The purpose of these theoretical results is to show that the general equilibrium determination of the optimal social security tax rate does not depend on the degree of heterogeneity in the productivity parameter φ. We can compute the socially optimal tax rate simply by maximizing the welfare of the most productive ($\varphi = 1$) cross-section of individuals which, by virtue of (28), ensures that the welfare of all other productivity cross sections will simultaneously be maximized.

4 Calibration and Welfare Results

We set calendar time $t = 0$, and we follow the related literature by using standard assumptions about parameter values. Labor-augmenting technology grows at rate $x = 1.56\%$ which follows the standard set in Bullard and Feigenbaum (2007). The rate of population growth is set to $n = 1\%$ to match the US experience in recent decades (İmrohoroğlu, İmrohoroğlu, and Joines 2003). The individual begins work at age 25, retires at age 65, and then dies at 80. Therefore, we set the length of the working period $T = 40$ and the economic lifespan $\bar{T} = 55$. Following convention and observed data, capital's share and the rate of depreciation are $\alpha = 35\%$ and $\delta = 8\%$. We normalize $A(0) = 1$ and $N(0) = 1$. Finally, the social security tax rate is set to $\theta = 10.6\%$ in order to match the OASI combined tax rate for employers and employees in the US.

This leaves only three free parameters: s_0, s_1, and φ_0. We set the lower bound on the saving rate distribution at $s_0 = 0$ to reflect the fact that some people save little or nothing for retirement. The remaining two parameters, φ_0 and s_1, can be chosen to ensure that the model generates a capital-output ratio equal to 2.5, following İmrohoroğlu, İmrohoroğlu, and Joines (2003). But we know from Lemma 1 above that φ_0 does not affect the capital-output ratio. Thus, the upper bound of the saving rate support determines the capital-output ratio. We find that $s_1 = 16\%$ gives $K(0)/Y(0) = 2.5$, with a corresponding equilibrium interest rate, r, equal to 6.0 percent.

About 20 percent of US households experience a discrete drop in consumption at the date of retirement (see Huang and Caliendo 2007 for a survey). Social security may potentially help those households who experience such a discrete drop by transferring income from the working years to the retirement years. For an individual born at time $\tau = 0$, a discrete drop in consumption occurs at the date of retirement (i.e., at $t = T$) if

$$(1-s)(1-\theta)\varphi w(T) > a(0|s,\varphi) + b(T|\varphi),$$

or

$$(1-s)(1-\theta)\varphi w(T) \quad > \quad \frac{\int_0^T s(1-\theta)\varphi w(t)e^{r(T-t)}dt}{\int_T^{\bar{T}} e^{r(T-t)}dt} + \frac{\theta w(T)L(T)\varphi}{\int_{T-\bar{T}}^0 \int_{\varphi_0}^1 \varphi f(\varphi)N(T)e^{n(\tau-T)}d\varphi \, d\tau}.$$

Now solve for the threshold saving rate

$$s < \frac{(1-\theta)w(T) - \dfrac{\theta w(T)L(T)}{\displaystyle\int_{T-\bar{T}}^{0}\int_{\varphi_0}^{1}\varphi f(\varphi)N(T)e^{n(\tau-T)}d\varphi\, d\tau}}{\left[(1-\theta)w(T) + \dfrac{\displaystyle\int_{0}^{T}(1-\theta)w(t)e^{r(T-t)}dt}{\displaystyle\int_{T}^{\bar{T}}e^{r(T-t)}dt}\right]}.$$

The threshold saving rate in our baseline calibration is 4.7 percent.[7] Since saving rates are distributed uniformly from 0 to 16 percent to match the target capital-output ratio, 29 percent of the population falls below the threshold and experiences a discrete drop in life-cycle consumption at the date of retirement. This is roughly in line with the data (20 percent, Huang and Caliendo 2007), which implies that the aggregate degree of "undersaving" in the model is realistic. Additionally, the average saving rate at our baseline calibration (8 percent) is close to the average saving rate among employer-sponsored plans (7 percent, Choi, Laibson, Madrian, and Metrick 2006). Hence, our model seems like a suitable instrument for studying the optimal provision of social security because it is consistent with reality along the dimensions relating to the aggregate capital stock, saving rates, and preparedness for retirement. Table 1 summarizes the baseline selection of parameters and the fit to the relevant data.

Table 2 reports the optimal tax rates at the baseline calibration (with $s_1 = 16\%$), for various parameterizations of the social welfare function (ρ, σ). Note that Proposition 1 ensures that the optimal tax rate is not a function of φ_0, so we do not need to specify this parameter. We compute in general equilibrium the optimal tax rate for each parameterization from a rectangle of standard values for the preference parameters, $\rho \in [0, 4\%]$, $\sigma \in [1, 5]$.[8] The optimal tax rate ranges from a low of 1 percent to a high of 13 percent. The average value of the optimal tax rate is 7.3 percent, which represents a modest program size relative

[7] Because $a(0|s, \varphi)$ and $b(T|\varphi)$ are both linear in φ, the threshold saving rate is invariant to φ. Thus, whether the individual experiences a discrete drop depends on his saving rate, but not his productivity.

[8] The parameter values of the welfare function do not affect the rule-of-thumb behavior at the micro level, nor do they affect the aggregate equilibrium. Hence, each possible parameterization in Table 2 shares the same calibration from Table 1.

to the current 10.6 percent tax rate in the US. However, if we restrict our attention to a zero discount rate, which is what many economists and commentators argue that a social planner ought to do (e.g., see Ramsey 1928 and the compilation of papers in Portney and Weyant 1999), then only the first row of Table 2 is relevant. The average of the first row is 10.2 percent, which is almost the same as the current rate in the US. Clearly, an increase in the discount rate lowers the optimal tax rate because the welfare gains from social security benefits are increasingly discounted relative to the welfare costs of taxes paid.

Table 3 reports sensitivity analysis of the optimal tax rate to the upper end of the saving distribution, s_1. Again, all calculations are in general equilibrium. However, as we adjust the upper bound, the model-generated endogenous variables (capital-output ratio, average saving rate, percentage who experience a discrete drop at retirement) will not remain at the baseline values found in Table 1. The discount rate is set to zero for each of these calculations, and the lower bound saving rate, s_0, is also fixed at zero. As we can see from Table 3, the optimal tax rate is not quantitatively sensitive to reasonable changes in the upper end of the saving distribution. It takes extreme changes to cause the optimal tax rate to change appreciably. Hence, the average optimal tax rate continues to take a value almost exactly the same as the size of the current program in the US for a modest range of possible saving distributions. There are general equilibrium forces at work that help the optimal tax rate remain nearly constant in s_1. On the one hand, the need for social security is weakened as the upper bound saving rate increases. But on the other hand, aggregate capital accumulation increases and the interest rate falls, which makes social security more attractive. It turns out that these two forces are quantitatively balanced for the moderate range of upper boundaries that we consider, $s_1 \in [12\%, 20\%]$.

It is important to emphasize that our model is biased against social security in at least one important way. Although our general equilibrium model does indeed make realistic predictions about the fraction of the population who will experience a discrete drop in consumption spending at the date of retirement, the average life-cycle consumption profile of a given cohort usually involves a discrete *jump* at the date of retirement. The high savers have a discrete jump while the lower savers have a discrete drop, and in general equilibrium it turns out that the average consumption profile for the cohort is dominated by the high savers. Thus, our baseline calibration involves oversaving on average, not undersaving. An equilibrium in which the average cohort consumption profile also displayed a discrete drop would involve a greater level of underpreparedness for retirement than is seen in our calibrations and hence would call for an even larger social security tax rate.[9]

This study is part of our broader research agenda which seeks to understand the optimal provision of social security under bounded rationality (e.g., Findley and Caliendo 2008, Findley and Caliendo 2009, Caliendo and Gahramanov 2009). In a companion paper in progress, we are trying to calibrate the aggregate life-cycle consumption profile of a given cohort to the hump-shaped consumption profile found in Gourinchas and Parker (2002) as well as the discrete drop in consumption spending at retirement. This exercise is straightforward in partial equilibrium. We would simply make the realistic assumption that wage income follows a hump shape over the life cycle (Gourinchas and Parker 2002), which in turns implies that consumption will also follow a hump shape, and then the interest rate and saving rate can be chosen in tandem to ensure a discrete drop at the date of retirement. However, the interest rate is endogenous in general equilibrium and we have not yet had success calibrating our rule-of-thumb model to both the hump and the drop. We leave this task for future work.

5 A Concluding Caveat

While this paper theoretically supports a pay-as-you-go program of the size in the US, the model has a number of important limitations that should prevent a focused political interpretation. For example, the optimal tax rates are computed under the assumption that the planner can choose between no social security program and a pay-as-you-go program of the variety studied here. We have not attempted to rank the many possible types of public pension arrangements (e.g., fully funded, partially funded, etc.), nor have we studied the many private saving incentives (e.g., IRA, 401k, SMarT) that a benevolent social planner may design to optimize lifetime utility. Hence, strictly speaking our results only rationalize

[9] Another assumption that biases the results against social security is that everyone in the model is a saver (with the exception of a zero-measure individual at the lower bound $s_0 = 0$). In reality, there is a large mass of people who save nothing for retirement.

a social security program relative to no program at all and are silent about the costs and benefits of proposals such as privatization.

Table 1. Baseline Calibration

Exogenous Parameter	*Symbol*	*Value*
Growth rate of technology	x	1.56%
Population growth rate	n	1%
Length of worklife	T	40 years
Length of economic lifespan	$\bar{T}$	55 years
Capital's share in income	α	35%
Depreciation rate	δ	8%
Technology size	$A(0)$	1
Size of cohort born at 0	$N(0)$	1
Social security tax rate	θ	10.6%
Lower bound saving rate	s_0	0
Upper bound saving rate	s_1	16%
Lower bound productivity	φ_0	irrelevant*

Target Variable	*Real World*	*Model Value*
Capital-output ratio	2.5	2.5

Verification of Calibration	*Real World*	*Model Value*
Average saving rate	7%	8%
% who experience consumption drop	20%	29%

*See Lemma 1 above.

Table 2. Baseline Calculations of the Optimal Social Security Tax Rate

		σ					
		1	2	3	4	5	avg.
	0	13%	9%	9%	10%	10%	**10.2%**
	1%	8%	8%	8%	9%	9%	**8.4%**
ρ	2%	5%	6%	7%	8%	9%	**7.0%**
	3%	3%	5%	6%	7%	8%	**5.8%**
	4%	1%	4%	6%	7%	8%	**5.2%**
	avg.	**6.0%**	**6.4%**	**7.2%**	**8.2%**	**8.8%**	**7.3%**

Note: The planner's discount rate is ρ and σ is the inverse elasticity of intertemporal substitution. The upper bound saving rate ($s_1 = 16\%$) is able to reproduce a realistic capital-output ratio, a realistic fraction of the population who experience a discrete drop in consumption at retirement, and a realistic mean saving rate. All calculations are in general equilibrium.

Table 3. Sensitivity Analysis of the Optimal Social Security Tax Rate to the Distribution of Saving Rates

		σ					
		1	2	3	4	5	avg.
	12%	13%	9%	9%	10%	10%	**10.2**%
	14%	13%	9%	9%	10%	10%	**10.2**%
s_1	16%	13%	9%	9%	10%	10%	**10.2**%
	18%	13%	9%	9%	10%	10%	**10.2**%
	20%	13%	9%	9%	10%	10%	**10.2**%
	avg.	**13.0%**	**9.0%**	**9.0%**	**10.0%**	**10.0%**	**10.2**%

Note: The social discount rate is set to $\rho = 0$. The upper bound saving rate is s_1 and σ is the inverse elasticity of intertemporal subsitution. All calculations are in general equilibrium.

References

[1] Ameriks, John, Andrew Caplin, and John Leahy (2003), Wealth Accumulation and the Propensity to Plan. *Quarterly Journal of Economics* 118(3), 1007-1047.

[2] Bucciol, Alessandro (2008), Social Security, Dynamic Efficiency and Self-Control Problems. Working paper, University of Amsterdam.

[3] Bullard, James and James Feigenbaum (2007), A Leisurely Reading of the Life-cycle Consumption Data. *Journal of Monetary Economics* 54: 2305-2320.

[4] Caliendo, Frank N. and Emin Gahramanov (2009), Hunting the Unobservables for Optimal Social Security: A General Equilibrium Approach. *Public Finance Review* 37(4), 470-502.

[5] Choi, James, David Laibson, Brigitte C. Madrian, and Andrew Metrick (2006), Saving for Retirement on the Path of Least Resistance. In Edward J. McCaffrey and Joel Slemrod, editors, Behavioral Public Finance: Toward a New Agenda, 304-351. New York: Russell Sage Foundation.

[6] Cremer, Helmuth, Philippe De Donder, Darío Maldonado, and Pierre Pestieau (2008), Designing a Linear Pension Scheme with Forced Savings and Wage Heterogeneity. *International Tax and Public Finance* 15(5): 547-562

[7] Docquier, Frédéric (2002), On the Optimality of Public Pensions in and Economy with Life-Cyclers and Myopes. *Journal of Economic Behavior and Organization* 47, 121-140.

[8] Feldstein, Martin (1985), The Optimal Level of Social Security Benefits. *Quarterly Journal of Economics* 100(2), 303-320.

[9] Findley, T. Scott and Frank N. Caliendo (2008), The Behavioral Justification for Public Pensions: A Survey. *Journal of Economics and Finance* 32(4), 409-425.

[10] Findley, T. Scott and Frank N. Caliendo (2009), Short Horizons, Time Inconsistency, and Optimal Social Security. *International Tax and Public Finance* 16(4), 487-513.

[11] Gourinchas, Pierre-Olivier and Jonathan A. Parker (2002), Consumption Over the Life Cycle. *Econometrica* 70, 47-89.

[12] Gul, Faruk and Wolfgang Pesendorfer (2001), Temptation and Self-Control. *Econometrica* 69(6), 1403-1435.

[13] Huang, Kevin X.D. and Frank N. Caliendo (2007), Rationalizing Seven Consumption-Saving Puzzles in a Unified Framework. Working paper, Vanderbilt University.

[14] İmrohoroğlu, Ayse, Selahattin İmrohoroğlu, and Douglas H. Joines (2003), Time Inconsistent Preferences and Social Security. *Quarterly Journal of Economics* 118(2), 745-784.

[15] Kotlikoff, Laurence J., Avia Spivak and Lawrence H. Summers (1982), The Adequacy of Savings. *American Economic Review* 72(5), 1056-1069.

[16] Kumru, Çağri S. and Athanasios C. Thanopoulos (2008), Social Security and Self Control Preferences. *Journal of Economic Dynamics and Control* 32(3), 757-778.

[17] Laibson, David (1997), Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics* 62: 443-477.

[18] Lusardi, Annamaria (2003), Planning and Saving for Retirement. Working Paper, Dartmouth College.

[19] Portney, Paul R. and John P. Weyant ed. (1999), Discounting and Intergenerational Equity (Resources for the Future: Washington, DC).

[20] Ramsey, Frank P (1928), A Mathematical Theory of Saving. *Economic Journal* 38: 543-559.

[21] Thaler, Richard H. (1994), Psychology and Savings Policies. *American Economic Review* 84(2), 186-192.

[22] Yakoboski, Paul and Jennifer Dickemper (1997), Increased Saving but Little Planning: Results of the 1997 Retirement Confidence Survey. *EBRI Issue Brief*, 191.