

# *Roth versus traditional accounts in a life-cycle model with tax risk\**

MARIE-EVE LACHANCE

*Department of Finance, College of Business Administration, San Diego State University,  
5500 Campanile Drive, San Diego, CA 92182-8236, USA  
(e-mail: Marie.Lachance@sdsu.edu)*

---

## Abstract

This paper analytically solves a life-cycle model that compares traditional and Roth retirement accounts. It includes realistic features such as tax deductibility of contributions and taxation of withdrawals, tax bracket structure with deductions, taxation of Social Security benefits, and tax risk at retirement. With current taxes, choosing a traditional account over a Roth creates small welfare losses in only a few cases, largely for those with higher incomes and pensions who are subject to the taxation of Social Security benefits. We also investigate tax variability and find that diversified strategies offer only small risk reduction benefits in our illustrations.

*Keywords:* Life-cycle models, Roth accounts, tax risk.

---

## 1 Introduction

A number of recent developments have sparked renewed interest in Roth retirement accounts. First, increasing government deficits can lead to higher tax rates in the future, reducing the appeal of traditional tax-deferred Individual Retirement Arrangements (IRAs) and 401(k)s. Second, new provisions in the tax code encourage the conversion of traditional accounts into Roth accounts.<sup>1</sup> Starting in 2010, the \$100,000 modified adjusted gross income (MAGI) limit for conversions has been removed and a special clause allows those who make the conversion in 2010 to split the proceeds equally between their 2011 and 2012 tax returns. Third, designated Roth accounts were introduced in the 401(k) arena in 2006. Although employer matching contributions are still restricted to traditional accounts, employees' contributions can be allocated to a Roth version of a 401(k) or 403(b).

This paper formally examines the differences between the front-loaded (traditional) and back-loaded (Roth) approaches to retirement accounts within the context of a

\*The author would like to thank an anonymous referee, Nikhil Varaiya, and Stefano Gubellini for helpful comments.

<sup>1</sup> See Dammon (2009) for an analysis of the conversion decision.

life-cycle model. Building on Yaari's (1965) classic framework with uncertain lifetime and borrowing constraints, the following realistic tax features are added: tax brackets and deductions, taxation of Social Security benefits, impact of retirement contributions/withdrawals on taxable income, and tax risk after retirement. This paper contributes to the literature by providing an analytical solution to the model for three cases: traditional account only, Roth account only, and any fixed combination of traditional and Roth accounts.

The dual approach proposed in this paper is a departure from more conventional dynamic programming techniques used to solve complex problems. This approach has three advantages: it handles better the discontinuity issues introduced by the realistic tax structure, it produces exact values, and it is fast. Working with accurate results is particularly important when examining small differences between two tax systems. The disadvantage of dynamic programming solutions is that they are based on unknown value functions, which must be estimated with time-consuming numerical optimization and interpolation. The broader methodological contribution of this paper is to show that the solution can be extracted instead from a set of known budget constraint equations, which eliminates the need to estimate unknown value functions and yields exact values. Further, the structure of the solution allows us to express the budget constraint with a series of closed-form equations. In terms of limitations, the model's main drawback is that it does not currently incorporate sources of background risk such as income risk, risky assets returns, or medical expenses risk.

The solution derived in this paper allows us to investigate three questions: Who benefits from Roth accounts? How does tax risk impact on the comparison? How does a policy of tax-deductible contributions impact on retirement savings? To answer these questions, we follow an approach commonly used in this literature and illustrate the model's solution with realistic income profiles for three education groups: less than high school, high school, and college. We also add a new dimension to the results by presenting them for three age cohorts (ages 25, 45, and 65 in 2010) and different levels of pension income.

When comparing traditional and Roth accounts, the standard argument is that traditional front-loaded accounts are a better option when marginal tax rates decline after retirement (see e.g. Engen *et al.* 1994).<sup>2</sup> The conventional wisdom is that this is generally the case and accordingly, traditional accounts have been favored historically. For example, Butterfield *et al.* (2000) conclude that 'traditional IRAs have significant wealth accumulation advantages over Roth IRAs in all but rare circumstances'. Our solution shows that this is not necessarily the case when the following tax loophole is taken into account: withdrawals from traditional accounts (but *not*

<sup>2</sup> Roth accounts have other advantages outside of this comparison. For instance, with traditional accounts the 10% penalty tax for early withdrawals applies to the entire withdrawal. For Roth accounts, the penalty is smaller as it applies only to the investment income portion. Roth accounts do not require minimum distributions starting at age 70½ and are perceived as better instruments to pass money tax-free to heirs. Another point raised in Burman *et al.* (2001) is that back-loaded options might be preferable when equal contributions limits apply because they shelter more money for retirement.

from Roth accounts) are considered when determining the amount of taxable Social Security benefits. For those affected, this implies that marginal tax rates after retirement effectively increase by either 50% or 85%.<sup>3</sup> Depending on the specific combination of tax rates, this either mitigates the tax decline after retirement or leads to an increase.

By contrasting expected utility in the pure Roth and traditional cases for the scenarios considered, we find that traditional accounts provide unambiguous benefits over their Roth counterparts mostly for those who pay no or little taxes after retirement.<sup>4</sup> For those at the other end of the spectrum with higher incomes and pensions, traditional accounts generate only small gains or losses due to the taxation of Social Security benefits. Currently, this issue affects mostly retirees in the group with a college degree, but the situation will be different for the younger generation as Social Security's taxation thresholds are not indexed for inflation.

Roth accounts may also become the preferred option if future tax rates increase. For instance, Kotlikoff *et al.* (2008) showed that they would be more appealing in a 30% tax hike scenario. We perform a different exercise by solving for the breakeven percentage increase in future tax rates required to make individuals indifferent between Roth and traditional accounts. For the group who pays no or minimal taxes after retirement, traditional accounts are still optimal even with a significant increase in taxes. For those who pay meaningful taxes, we find that the breakeven increases range from 4% to 41%. Although a modest increase in future taxes could tip the scale in favor of Roth accounts for those with high pensions, a more radical change would be required for those without pensions.

To address concerns about tax risk, diversified strategies that mix traditional and Roth accounts have been commonly suggested. Yet, little formal evidence has been provided to support them. Dickson (2004) is one of the rare models to include tax variability in a two-period model with both traditional and Roth accounts.<sup>5</sup> In this paper, we extend the concept of tax variability to a life-cycle framework, which allows us to measure the benefits of tax diversification in a realistic setup. We consider a simple naïve diversification strategy where, in every period, the individual allocates the same fraction of his savings/withdrawals to the traditional and Roth accounts. We solve for the optimal fraction to invest in traditional accounts in the cases with and without tax risk and find that the optimal allocation is essentially the same in both cases. In the scenario without tax risk, the option to invest a fraction of savings in Roth accounts can be worth a few thousand dollars because it can help avoid higher marginal tax rates. In contrast, risk reduction benefits have more limited

<sup>3</sup> These multipliers do not apply to those with either very low or very high incomes. Note that the idea that the taxation of Social Security benefits can increase lifetime taxes with 401(k)s was introduced in Gokhale *et al.* (2001). This paper's contribution is to incorporate this concept within the analytical solution of a life-cycle model.

<sup>4</sup> This is for the case with standard deductions and no tax penalty. Roth accounts could also be recommended for those who fall in a lower tax bracket because of especially high deductions or a temporary loss of income.

<sup>5</sup> He shows that Roth accounts are superior when tax rates do not decrease after retirement as they are less risky. As the spread between pre- and post-retirement tax rates increases, a greater proportion of savings should be allocated to traditional accounts.

impact on the allocation decision because they amount to a few hundred dollars at best.<sup>6</sup>

Finally, our model allows us to examine the differential impact of traditional accounts on retirement savings and consumption.<sup>7</sup> Our results indicate that, although traditional accounts can increase gross retirement savings substantially for those with a college degree (by up to \$68,000 in the baseline scenario), most of this increase vanishes once we take into account the present value of taxes that will have to be paid on withdrawals. In other words, for that group traditional accounts increase the size of assets under management, but not necessarily retirement consumption. Those in the less-than-high-school group present a different story as they do not pay taxes after retirement and actually display the highest increase in retirement consumption at \$15,000. Decomposing the increase into an income effect (from the tax subsidy) and a substitution effect (from increased savings before retirement), we find that the substitution effect has more impact on retirement consumption because a large fraction of the tax subsidy is used to increase consumption *before* retirement.

The remainder of this paper is structured as follows. Section 2 describes the model and Section 3 provides the solution. Section 4 lists the assumptions used in the numerical illustrations and Section 5 illustrates some representative cases. Section 6 analyzes who gains from Roth accounts and considers the potential benefits of mixed strategies. Section 7 details the impact of tax deductible contributions on the level of retirement savings. Section 8 concludes with some suggestions for applications and future research.

## 2 Model

The model builds on Yaari's (1965) classic life-cycle framework, which features borrowing constraints and an uncertain lifetime. The model's contribution to the life-cycle literature is to incorporate many realistic features of the tax treatment of retirement savings, while maintaining an analytical structure. First, the model can be used to model either Roth or traditional accounts. In the Roth case, contributions and withdrawals do not affect taxes. In the traditional case, contributions are deductible from taxable income and withdrawals are taxed. The accounts can be part of an IRA or a 401(k). Second, while many models use a single fixed tax rate, this model reflects the United States tax structure with various tax brackets and deductions. Third, the model lets tax rates be endogenously determined rather than being specified exogenously. Fourth, the model incorporates the actual rules for the taxation of Social Security benefits. Fifth, the model allows for tax risk after retirement.

Although it would also be interesting to let the individual choose between Roth and traditional accounts in each period, this setup would make the problem more

<sup>6</sup> In the scenario without risk, those who gain from a mixed strategy are those who are subject to higher marginal tax rates after retirement and are able to reduce these rates by taking smaller withdrawals from traditional accounts.

<sup>7</sup> The previous literature (e.g. Engen *et al.* 1994; Laibson *et al.* 1998; Gomes *et al.* 2009) has examined the impact of traditional accounts over taxable accounts. Our results are not directly comparable as traditional accounts are benchmarked against Roth accounts instead of taxable accounts, effectively measuring the differential impact of the tax deductibility of contributions.

cumbersome to solve. The model can be used, however, to offer a solution for the simpler case of a naïve diversification approach. With that strategy, the proportions allocated to the traditional and Roth accounts are, respectively,  $\alpha$  and  $1 - \alpha$  in every period. This solution allows us to illustrate the potential value of diversification in a context with tax risk. In the remainder of the text, all equations are given as a function of  $\alpha$  since the pure cases can be viewed as special cases as follows:

$$\begin{aligned} \alpha = 1 & \quad \text{Traditional account only.} \\ \alpha \in (0, 1) & \quad \text{Fraction } \alpha \text{ in traditional account and } 1 - \alpha \text{ in Roth account.} \\ \alpha = 0 & \quad \text{Roth account only.} \end{aligned} \quad (1)$$

To balance the increased complexity brought by realistic taxes, straightforward assumptions are used for the rest of the model. This approach has the advantage of leading to a solution with exact values which lends itself well to analysis. For the elements not included in this model, the reader is referred to: Cocco *et al.* (2005) (income risk and portfolio choice), Love (2007) (employer contributions), and Kotlikoff *et al.* (2008) (married individuals).

### 2.1 Economic and demographic assumptions

The individual is age  $t_0$  when the problem starts and he can live up to age  $\omega$ . The probability that he survives from age  $t_0$  to age  $t$  is denoted by  $p_{t_0, t}$ . It is assumed that  $p_{t_0, t}$  is continuous, decreases with time, and eventually converges to zero as  $t \rightarrow \omega$ . The utility of consumption is represented by an increasing and concave function  $u(c)$  with  $u'(c) > 0$ ,  $u''(c) < 0$ ,  $\lim_{c \rightarrow 0} u'(c) = \infty$ , and  $\lim_{c \rightarrow \infty} u'(c) = 0$ . Time preferences are taken into account by discounting utility at a continuous rate  $\beta$ . The combined discount from time and mortality is denoted by

$$f(t) = e^{-\beta(t-t_0)} p_{t_0, t}. \quad (2)$$

All economic assumptions are expressed in real terms. Savings grow at a real risk-free rate  $r > 0$ . The wealth process is denoted by  $W_t$  and it is assumed that the individual starts the problem with no initial savings, i.e.  $W_{t_0} = 0$ . Borrowing is not allowed. Before retirement, pre-tax income is a continuous function denoted by  $y_t$ . The individual retires at age  $R$ , which is assumed to be 65 years old. After retirement, the individual receives a Social Security pension with annual payments of  $SS > 0$ . He may also receive annual income  $y_R \geq 0$  from an employer pension or another source of annuity payments, for a total of  $y_t = y_R + SS$  per year. The solution considers only the case where expected after-tax income declines at the time of retirement.

### 2.2 Taxation

Before retirement, income is subject to a payroll tax rate  $\pi$ . At all times, federal income taxes based on the United States tax bracket system apply.<sup>8</sup> For  $k = 0, \dots, K$ ,

<sup>8</sup> The model does not currently include State taxes because of the large number of cases to consider. For those who pay taxes before retirement only, including State taxes would imply higher tax subsidies and a greater incentive to save for retirement. The impact would be more limited for those who pay taxes after retirement if marginal tax rates increase by the same amount before and after retirement.

the brackets are denoted by  $[B_k, B_{k+1})$  and the marginal tax rate within bracket  $k$  is  $\tau_k$ . Tax rates after retirement are subject to a one-time multiplicative shock  $\theta$  that applies to all marginal tax rates. To represent this risk, the model allows choosing across the entire family of discrete probability distribution. Let  $N$  be the number of possible states, there is a probability  $p_i$  with  $i=1, \dots, N$  that  $\theta=\theta_i$  and  $\tau_k^i = \theta_i \tau_k$ . The case without tax risk can be viewed as a special case of this model with  $N=1$  and  $\theta_1=1$ . For the remainder of this paper, all functions of  $\tau_k$  should be interpreted as random variables after retirement although the notation is not differentiated.

To compute taxable income, the standard deduction and personal exemption must be taken into account. The sum of these two components is denoted by  $E$  before age 65 and by  $E_R$  after age 65 (there is an increase in the standard deduction at age 65). In the traditional case, it must also be recognized that taxable income is reduced by contributions before retirement and increased by withdrawals after retirement. The contributions or savings before retirement are denoted by  $s_t$  where  $s_t = y_t - c_t - \text{tax}_t(s_t)$  and  $c_t$  is consumption. A negative  $s_t$  represents a withdrawal. Let  $\text{SS}_t^{\text{tx}}$  denote the taxable portion of Social Security benefits, the taxable income function is given by

$$y_t^{\text{tx}}(s_t) = \begin{cases} y_t - \alpha s_t - E, & t < R, \\ y_R + \text{SS}_t^{\text{tx}}(s_t) - \alpha s_t - E_R, & t \geq R. \end{cases} \tag{3}$$

Accordingly, for  $k=0, \dots, K$ , the tax function is

$$\text{tax}_t(s_t) = \pi y_t \cdot 1(t < R) + \tau_k y_t^{\text{tx}}(s_t) + G_k, \quad B_k \leq y_t^{\text{tx}}(s_t) < B_{k+1}, \tag{4}$$

where

$$G_k = \sum_{j=1}^{k-1} \tau_j (B_{j+1} - B_j) - \tau_k B_k. \tag{5}$$

It should be noted that some of the features of the tax system are not included in the model because they do not affect the solution in most cases illustrated here. For instance, the model does not limit the amount of tax deductible contributions because the optimal savings in all our illustrations are below the 401(k)'s \$16,500 limit. In addition, the 10% penalty tax for early withdrawals before age  $59\frac{1}{2}$  is not incorporated because the model does not feature shocks that would trigger them. This eventuality is discussed later in Section 6.3.

The model is completed by specifying  $\text{SS}_t^{\text{tx}}(s_t)$ , the taxable portion of Social Security benefits. This is a fairly complex formula. First, a provisional income measure  $\text{PI}_t \equiv \text{PI}(s_t)$  is defined by adding half of Social Security benefits to other sources of income (including withdrawals from traditional accounts, but not Roth accounts), yielding:

$$\text{Provisional Income}_t = \text{PI}(s_t) = y_R + \frac{\text{SS}}{2} - \alpha s_t, \quad t \geq R. \tag{6}$$

Once  $\text{PI}_t$  is computed, it has to be compared to two 'base amounts'  $X^1$  and  $X^2$ . Currently, these base amounts are  $X^1 = \$25,000$  and  $X^2 = \$34,000$  for singles; the corresponding measures are \$32,000 and \$44,000 for those married filing jointly.

These amounts are not indexed for inflation. Given that the rest of the model is expressed in real terms, these thresholds have to be adjusted for inflation (denoted by  $i$ ) and this creates cohort effects in the model. For example, for somebody retiring at age 65 in 2010, the first base amount in real terms is \$25,000. For somebody currently age 45 in 2010, this same base amount will be \$13,720 when they retire in 20 years (assuming  $i = 3\%$ ), making it more likely that they will be taxed. For a given individual, the base amounts will also decrease by  $e^{-it}$  in real terms after  $t$  years of retirement. Thus, let  $a$  be the current age in 2010, the base amounts in real terms can be expressed by

$$X_t^1 = X^1 e^{-i(t-a)} \quad \text{and} \quad X_t^2 = X^2 e^{-i(t-a)}. \tag{7}$$

If the provisional income  $PI_t$  is less than the first base amount  $X_t^1$ , none of the Social Security benefits are taxable. If  $PI_t$  is greater than the first base amount, but lower than the second, 50% of the excess  $PI_t - X_t^1$  is taxable income.<sup>9</sup> If  $PI_t$  is greater than the second base amount, then the taxable portion is  $50\%(X_t^2 - X_t^1) + 85\%(PI_t - X_t^2)$ , subject to a maximum of 85% of Social Security benefits. Thus, depending on the level of  $PI_t$ , there are four different formulas that can apply: 0,  $50\%(PI_t - X_t^1)$ ,  $50\%(X_t^2 - X_t^1) + 85\%(PI_t - X_t^2)$ , and 85% SS. To summarize, these four cases can be embedded into one linear structure:

$$SS_t^{tx}(s_t) = M_h PI(s_t) + H_{t,h}, \quad B_{t,h}^S \leq PI(s_t) < B_{t,h+1}^S, \tag{8}$$

where

$h$	$M_h$	$H_{t,h}$	$B_{t,h}^S$
0	0	0	0
1	50%	$-50\% X_t^1$	$X_t^1$
2	85%	$-50\% X_t^1 - 35\% X_t^2$	$X_t^2$
3	0	$85\% \cdot SS$	$SS + \frac{35\%}{85\%} X_t^2 + \frac{50\%}{85\%} X_t^1$

The key notation to remember for the rest of the analysis is  $M_h$ , which represents the marginal rate of inclusion of Social Security benefits in taxable income. From equations (6) and (8), a dollar withdrawn from a traditional account increases the taxable portion of Social Security benefits by  $M_h$  and triggers an additional tax of  $\tau_k M_h$ . Thus,  $M_h$  can be viewed as a factor that magnifies the marginal tax rate from  $\tau_k$  to  $\tau_k(1 + M_h)$ . For the remainder of this paper,  $\tau_k(1 + M_h)$  will be referred to as the effective marginal tax rate. For example, if  $\tau_k = 15\%$  and  $M_h = 85$ , withdrawals from traditional accounts are effectively taxed at a rate of  $15\% (1 + 85\%) = 27.75\%$ .

To put this in perspective, recall that the basis for favoring traditional accounts is that marginal tax rates after retirement should be lower than before retirement. Once the taxation of Social Security benefits is taken into account, this logic may no longer hold as it is possible to have lower income after retirement, but higher tax rates. This issue does not necessarily affect everyone, those with either very low or very high

<sup>9</sup> There is a maximum of 50% of Social Security benefits that apply to this taxable amount. This rule is not incorporated in the model.

incomes are not impacted because their inclusion rate  $M_h$  is zero. It should be kept in mind that the composition of this group varies across cohorts because the base amounts change in real terms. Initially, those with higher incomes are subject to the 50% or 85% inclusion rates. Eventually, they will see their  $M_h$  decline to zero when their taxable benefits reach 85% SS. On the other hand, those with lower incomes start with  $M_h=0$  but ultimately see their inclusion rates increase to 50% and 85% as the base amounts decrease in real terms.

### 2.3 Optimization problem

Combining these assumptions, the individual's problem is to choose consumption to maximize his expected utility, subject to a budget constraint and borrowing constraints. Thus, the optimization problem is

$$\max_{c > 0} E \left[ \int_{t_0}^{\omega} f(t)u(c_t)dt \right] \tag{10}$$

such that

$$dW_t = [W_{t,r} + y_t - c_t - \text{tax}_t(s_t(c_t))]dt \tag{11}$$

and

$$W_t \geq 0 \quad \forall t, \quad W_{t_0} = 0, \tag{12}$$

where  $\text{tax}_t(s_t)$  is given in (4) and  $s_t(c_t)$  by

$$s_t(c_t) = \frac{y_t - \text{tax}_t(0) - c_t}{1 - \alpha\tau_k(1 + M_h)} \tag{13}$$

for  $B_k \leq y_t^{\text{tx}}(s_t) < B_{k+1}$  and  $B_{t,h}^S \leq \text{PI}(s_t) < B_{t,h+1}^S$  with  $y_t^{\text{tx}}(s_t)$  given in (3) and  $\text{PI}(s_t)$  given in (6).

### 2.4 Discontinuity points for $\tau_k$ and $M_h$

The rates  $\tau_k$  and  $M_h$  in this model are step functions of consumption and these discontinuities affect the structure of the solution in the next section. To facilitate the exposition of the solution, the notations  $C_k$  and  $C_h$  are introduced to identify the levels of consumption where respectively the marginal tax rate jumps from  $\tau_{k-1}$  to  $\tau_k$  and the rate of inclusion of Social Security benefits jumps from  $M_{h-1}$  to  $M_h$ . We can solve for  $C_k$  by setting taxable income equal to the next tax bracket (i.e.  $y_t^{\text{tx}}(s(C_k)) = B_k$ ) and inverting  $C_k$  from equation (13) as follows:<sup>10</sup>

$$C_k = \begin{cases} y_t(1 - \pi - 1/\alpha) + B_k(1/\alpha - \tau_k) - G_k + E/\alpha, & t < R, \\ y_R(1 - \frac{1}{\alpha}) + \text{SS} \left( 1 - \frac{M_h/2}{\alpha(1 + M_h)} \right) + B_k \left( \frac{1}{\alpha(1 + M_h)} - \tau_k \right) - G_k + \frac{-H_{t,h} + E_R}{\alpha(1 + M_h)}, & t \geq R. \end{cases} \tag{14}$$

<sup>10</sup> In (14),  $h$  is such that  $C_k \in (C_h, C_{h+1})$ ; in (15),  $k$  is such that  $y_t^{\text{tx}}(C_h) \in [B_k, B_{k+1})$ .  $C_k$  and  $C_h$  are defined only when positive.

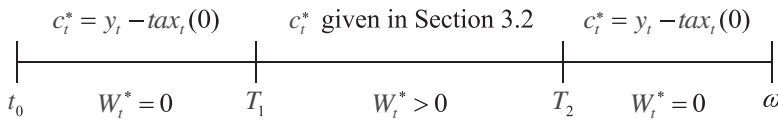


Similarly,  $C_h$  is derived by setting the provisional income in equation (6) equal to the next threshold  $B_{t,h}^S$  to obtain:

$$C_h = y_R \left( 1 - \frac{1}{\alpha} \right) + SS \left( 1 + \frac{\tau_k \alpha - 1}{2\alpha} \right) + B_{t,h}^S \times \left( \frac{1}{\alpha} - \tau_k(1 + M_h) \right) - G_k - \tau_k(H_{t,h} - E^R). \tag{15}$$

### 3 Solution

We derive a complete solution and proof for the optimization problem in (10)–(13). This section outlines the solution’s key components and we refer to Appendix A for additional technical details. Readers less interested in the derivation of the results can proceed to the illustrations in Section 4. In a problem with borrowing constraints, the binding periods with  $W_t^* = 0$  must be identified because the solution is different in these periods as the individual simply consumes all his after-tax income. To simplify the exposition in this section, we consider the case where the individual starts saving for retirement at age  $T_1$  and exhausts his savings by age  $T_2$ , i.e. we have the following structure:



#### 3.1 Lagrangian and dual approach

Solving the problem with conventional methods can be challenging because we need to prove that the borrowing constraint is satisfied everywhere. Without restricting the model’s assumptions, this is a difficult task because the optimal consumption function provides little insight into the sign of the wealth process. In this context, it is useful to turn to the dual approach suggested in Lachance (2012) and extend it to include discontinuities and risks. Dual approaches are used with problems that are difficult to solve in their primal form, but are easier to handle in their equivalent dual form. In this case, the advantage is that we do not need to prove directly that  $W_t^* \geq 0$  everywhere and this allows us to get a simpler condition.

As with conventional optimization methods, to prove that a solution is optimal the problem’s Lagrangian must be derived first. Appendix A.1 describes how the standard Lagrangian is constructed and rewrites it in a form suitable for the dual approach as follows:

$$L = E \left[ \int_{t_0}^{\omega} (f(t)u(c_t) + X(t)e^{-r(t-t_0)}s_t(c_t))dt \right]. \tag{16}$$

The process  $X(t)$  is just a transformation of the original Lagrange multipliers; it must be decreasing during the binding periods and equal to a constant  $\lambda$  during the

non-binding period  $[T_1, T_2]$ . With tax risk after retirement, a different constant  $\lambda_i$  applies in each state  $i$ . By construction, the constants are related by  $\lambda = \sum_{i=1}^N p_i \lambda_i$  and can be interpreted as the marginal utility of wealth. Less technically, these conditions on  $X(t)$  mean that expected utility cannot be improved by saving an additional dollar at time  $t$  and spending it later (and vice-versa).

The first step of the dual approach is to find  $c_t^*$  that maximizes  $L$ , which is a simple unconstrained maximization problem detailed in Section 3.2. The second step of the dual approach is to find  $X(t)$  that minimizes  $L$ , which in our setup is equivalent to:

- Solving for the constants  $\lambda$  and  $\lambda_i$  that satisfy the budget constraint equations  $W(\lambda, \lambda_i) = 0, i = 1, \dots, N$  and the condition  $\lambda = \sum_{i=1}^N p_i \lambda_i$ .
- If there is more than one constant  $\lambda$  or  $\lambda_i$  that satisfies the budget constraint, the highest one is optimal.

A solution that satisfies these criteria maximizes utility and satisfies all the problem's constraints. Appendix A.2 details the equation for the budget constraint  $W(\lambda, \lambda_i) = 0$  and Appendix A.3 explains how the bisection method can be employed to solve for the constants  $\lambda$  and  $\lambda_i$ .

### 3.2 Optimal consumption and issues with discontinuities

To find  $c_t^*$  that maximizes  $L$  unconditionally when  $X(t) = \lambda$ , the standard procedure is to derive the first-order condition (F.O.C.) as follows:

$$L'(c_t) = f(t)u'(c_t) - \frac{\lambda e^{-r(t-t_0)}}{1 - \alpha\tau_k(1 + M_h)}, \tag{17}$$

$$\text{F.O.C.: } L'(c_t^*) = 0 \Rightarrow c_t^*(\lambda) = u'^{-1}\left(\frac{\lambda e^{-r(t-t_0)}}{f(t)(1 - \alpha\tau_k(1 + M_h))}\right).$$

The first component in  $L'(c_t)$  measures the marginal benefit of consumption and the second one the opportunity cost of saving where  $s_t'(c_t) = -1/(1 - \alpha\tau_k(1 + M_h))$ . With traditional accounts, savings do not decrease one-for-one with consumption because taxes create an additional reward (or penalty for withdrawals). In that case, higher taxes translate into higher savings before retirement and lower withdrawals after retirement.

When  $L'(c_t)$  is continuous,  $L''(c_t) < 0$  and there is a unique solution to the F.O.C.  $L'(c_t^*) = 0$ . When  $\tau_k$  and  $M_h$  enter  $L'(c_t)$ , this cannot be taken as granted because the discontinuities raise new technical issues: (1) the F.O.C. may not have a solution  $c_t^*$  for every  $\lambda$ , (2) the F.O.C. can have more than one solution  $c_t^*$  for a given  $\lambda$ , and (3) circularity, i.e.  $\tau_k$  is a function of  $c_t^*$  and  $c_t^*$  is a function of  $\tau_k$ . Note that these discontinuity issues can also arise in retirement problems where retirement savings are rewarded (or withdrawals penalized) and the reward/penalty can suddenly change. For example, this would be the case if an employer offers a 50% match on employee contributions, but stops the match for contributions above 6%.

For the first issue, recall that Section 2.4 defined the discontinuity points  $C_k$  and  $C_h$  where  $\tau_k$  and  $M_h$  jump. Here, the notation  $C$  will stand for any of these points. The problem is that it is not possible to solve for  $c$  in  $L'(c) = 0$  when at a discontinuity

point  $C$  we have  $L'(C^-) > 0$  and  $L'(C) < 0$ . Actually, with  $L''(c) < 0$ ,  $C$  is locally optimal because  $L(c)$  increases up to  $C$  and decreases thereafter. Since the equation for  $L'(C)$  in (17) is a function of  $\lambda$ , it will be convenient to rewrite the condition in terms of a range of values for  $\lambda$  as follows:

$$L'(C^-) > 0 \quad \text{and} \quad L'(C) < 0 \quad \text{is equivalent to} \quad \Lambda_t(C) < \lambda < \Lambda_t(C^-), \quad (18)$$

where the function

$$\Lambda_t(c) = u'(c) f(t) e^{r(t-t_0)} (1 - \alpha \tau_k (1 + M_h)) \quad (19)$$

is simply introduced to express the solution more compactly later on.<sup>11</sup> Outside these values for  $\lambda$ , the standard F.O.C. solution in (17) applies. This approach with brackets for  $\lambda$  offers a simple test to determine when an algorithm should use the standard F.O.C. solution in (17) and when it should use the discontinuity points  $C$ . It has also the advantage of addressing the circularity issue: for a given  $\lambda$ , we know which  $\tau_k$  or  $M_h$  to use in  $c_t^*(\lambda)$ .

The second problem that we can encounter is having more than one locally optimal solution. This happens when  $L'(c_t)$  increases at a discontinuity, i.e. in the uncommon scenario where the effective marginal tax rate decreases with consumption. In this model, this occurs at the point  $C_{h=3}$  where taxable Social Security benefits reach the 85% maximum and  $M_h$  drops from 85% to 0%. Specifically, when  $L'(C_{h=3}^-) < 0$  and  $L'(C_{h=3}) > 0$ , there are two locally optimal solutions  $c_1 < C_{h=3}$  and  $c_2 > C_{h=3}$  (with  $h < 3$  for  $c_1$  and  $h = 3$  for  $c_2$ ). The globally optimal one maximizes  $L(c_t)$  and to express this more systematically, we solve for a point  $\lambda = \bar{\Lambda}$  such that the individual is indifferent between  $c_1$  and  $c_2$ . It is possible to show that  $c_1$  is optimal for all  $\lambda > \bar{\Lambda}$ ,  $c_2$  is optimal for all  $\lambda < \bar{\Lambda}$ , and the solution jumps between  $c_1$  and  $c_2$  when  $\lambda = \bar{\Lambda}$ .<sup>12</sup> Figure 1 makes this result more intuitive by illustrating the solution graphically for the cases  $\lambda < \bar{\Lambda}$ ,  $\lambda = \bar{\Lambda}$ , and  $\lambda > \bar{\Lambda}$ .

Combining these results, the optimal solution in the non-binding period after retirement  $t \in [R, T_2)$  can be condensed with:

$$c_t^*(\lambda) = \begin{cases} u'^{-1} \left( \frac{\lambda e^{-r(t-t_0)}}{f(t)(1-\alpha\tau_k(1+M_h))} \right), & \max[\Lambda_t(C_{k+1}^-), \Lambda_t(C_{h+1}^-)] \leq \\ C_h, & \lambda \leq \min[\Lambda_t(C_k), \Lambda_t(C_h)], \\ C_k, & \Lambda_t(C_h) < \lambda < \Lambda_t(C_h^-), \quad h < 3, \\ & \Lambda_t(C_k) < \lambda < \Lambda_t(C_k^-), \end{cases} \quad (20)$$

for  $\lambda \geq \bar{\Lambda}$  and

$$c_t^*(\lambda) = \begin{cases} u'^{-1} \left( \frac{\lambda e^{-r(t-t_0)}}{f(t)(1-\alpha\tau_k)} \right), & \Lambda_t(C_{k+1}^-) \leq \lambda \leq \Lambda_t(C_k), \\ C_k, & \Lambda_t(C_k) < \lambda < \Lambda_t(C_k^-), \end{cases} \quad (21)$$

for  $\lambda < \bar{\Lambda}$ . Before retirement, the points  $C_h$  related to Social Security do not apply and the solution in the non-binding period  $[T_1, R)$  is given by (21) for all  $\lambda$ . In the pure

<sup>11</sup> In  $\Lambda_t(c)$ ,  $k$  is such that  $y_t^{\text{bx}}(st(c)) \in [B_k, B_{k+1})$  and  $h$  is such that  $\text{PI}(s_t(c)) \in [B_{t,h}^S, B_{t,h+1}^S)$ .

<sup>12</sup> Let  $\phi(\lambda) = L(c_1, \lambda) - L(c_2, \lambda)$  and  $\bar{\Lambda}$  be such that  $\phi(\bar{\Lambda}) = 0$ , since  $\phi'(\lambda) > 0$ ,  $c_1$  maximizes  $L$  when  $\lambda > \bar{\Lambda}$  and  $c_2$  maximizes  $L$  when  $\lambda < \bar{\Lambda}$ . For  $C = C_{h=3}$ , it can be shown that  $\phi(\Lambda(C^-)) < 0$  and  $\phi(\Lambda(C)) > 0$ , which guarantees the existence of a unique  $\bar{\Lambda} \in (\Lambda(C^-), \Lambda(C))$ .

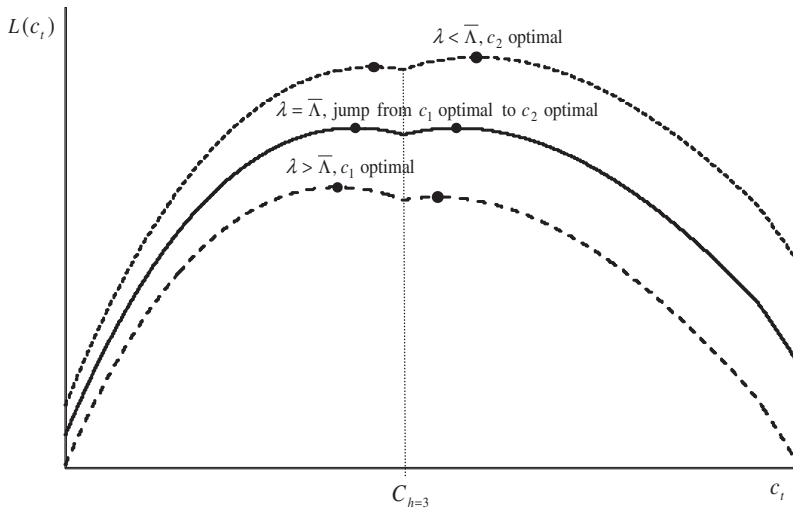


Figure 1. Illustration of jump in optimal consumption (optimal  $c_t$  maximizes  $L(c_t)$ ).

Roth case ( $\alpha=0$ ), only the top part of the solution applies. In other words, working with Roth accounts in optimization problems is much easier than working with traditional accounts.

If the utility function takes the standard power utility form  $u(c) = c^{1-\gamma}/(1-\gamma)$  with  $\gamma \neq 1$ , the top portion of the solution in (20) and (21) becomes:

$$c_t^*(\lambda) = \left( \frac{\lambda e^{(\beta-r)(t-t_0)}}{p_{t_0,t}(1-\alpha\tau_k(1+M_h))} \right)^{-1/\gamma} \tag{22}$$

Appendix B shows how this equation can be combined with a flexible set of assumption for mortality and income to express the budget constraint  $W(\lambda, \lambda_t) = 0$  as a series of closed-form equations.

### 3.3 Comparison with dynamic programming

In the previous literature, dynamic programming techniques have been the tool of choice for solving life-cycle models with risks and constraints, and it can be useful to contrast them with the approach described in this section. With dynamic programming, the problem is to solve the Bellman equation  $V(W_t) = \max_{c_t} [u(c_t) + \beta p_t E[V_{t+1}(W_{t+1})]]$  subject to the budget constraint in (11), which is essentially a discretized version of our earlier problem of maximizing  $L(c_t)$ . Since the functional form for the value function  $V_{t+1}(W_{t+1})$  is not known, it must be estimated numerically by backward induction for a set of values, and then interpolated between these points. With this interpolated function, the simplest approach to solve the Bellman equation is to test every possible value of  $c_t$ : this is a slow process and has limited precision unless a fine grid is used. The circularity issue mentioned previously also makes the process more time consuming.

Although several numerical techniques can be used to perform a more efficient search for optimal consumption, their application is not straightforward with discontinuities. For instance, most assume the existence of a single optimum or solution within a given interval, but Figure 1 illustrated that this premise can fail. Furthermore, techniques based on first-order-conditions or related Euler equations have to be modified because the solution in (20) and (21) showed that several cases have to be considered. In other words, applying these techniques to problems with discontinuities requires some adjustments and an analytical process such as the one in Section 3.2 can be followed to determine the nature of the changes.

The technique used in this paper is particularly interesting because the solution is based on a known budget constraint instead of an unknown value function, which eliminates the need for interpolation and leads to exact values. In addition, it is not necessary to numerically optimize  $c_t^*$  as the solution is given in (20) and (21). Arguably, the approach suggested here is limited in the sense that it is based on a one-time risk, but the same concepts could be extended to multiple risky periods. With  $M$  possible risky paths, the crux of the problem would be to solve a system of  $M$  equations (the budget constraints, possibly as a series of closed-form expressions) and  $M$  unknowns  $\lambda_1, \dots, \lambda_M$ .<sup>13</sup> If this system can be solved in a reasonable amount of time, then the approach can be an interesting alternative to dynamic programming. With a one-time risk, the problem proved to be particularly manageable as it can be reduced to a one-equation-in-one-unknown problem that can be solved with the bisection method.

#### 4 Assumptions for numerical illustrations

We use the equations detailed in Appendix B and the assumptions detailed in this section to generate the numerical illustrations in the remainder of this paper. The individual starts the problem at age  $t_0=25$  and lives up to age  $\omega=105$ . The model uses a discrete mortality table, which is calibrated with survival probabilities derived from the National Center for Health Statistics 2005 data.<sup>14</sup> The baseline parameters for respectively the power utility function, the rate of return, and the time discount factor are given by:  $\gamma=3$ ,  $r=3\%$ , and  $\beta=3\%$ .

As in Cocco *et al.* (2005), the income profiles are calibrated using data from the Panel Study of Income Dynamics (PSID) for three education levels: less than high school (LHS), high school (HS), and college.<sup>15</sup> Figure 2 illustrates the income profiles for each of the three education groups. Social Security benefits are computed

<sup>13</sup> While a formal extension to the case with multiple risky periods is left for future research, we can give the intuition here by adding one more risky period. For example, if after each state  $i$  there are  $J$  possible states  $j=1, \dots, J$ , there would be an additional set of  $N \cdot J$  constants  $\lambda_{i,j}$  and  $N$  conditions  $\lambda_i = \sum_{j=1}^J p_j \lambda_{i,j}$ . The budget constraint would take the form  $W(\lambda, \lambda_i, \lambda_{i,j})=0$  and the problem would reduce to a system of  $N \cdot J = M$  equations in  $M$  unknowns  $\lambda_{i,j}$  because  $\lambda$  and  $\lambda_i$  can be expressed in terms of  $\lambda_{i,j}$ .

<sup>14</sup> The survival probabilities are derived using standard actuarial techniques and adding the number of deaths for males and females to obtain unisex rates.

<sup>15</sup> Those with some college but no degree are not included in these categories. The estimation includes PSID waves from 1970 to 2007. Results are based on median values for the household head's labor income variable, with an adjustment to 2010 dollars using the Consumer Price Index for urban wage earners and clerical workers. The data includes households headed by both males and females whose employment status is either working now, temporarily laid-off, or unemployed looking for work.

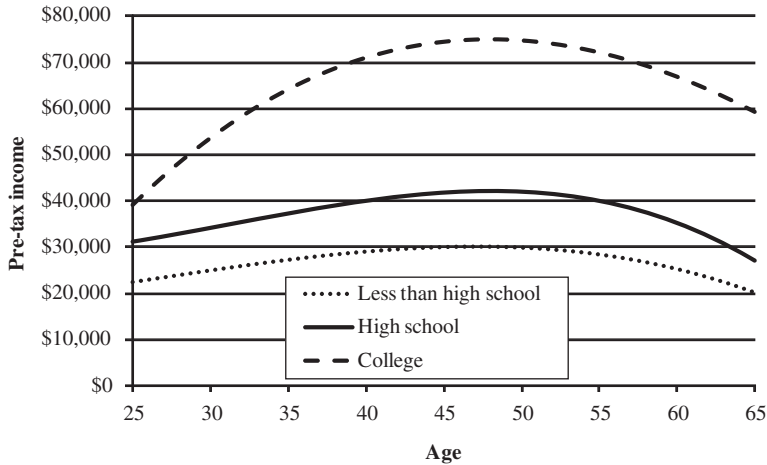


Figure 2. Income profiles by education.

according to the formula applicable in 2010.<sup>16</sup> The annual benefits are respectively \$13,239 for the less-than-high-school group, \$16,481 for the high-school group, and \$23,278 for those with a college degree. For the pension income assumption, the use of a baseline ‘average’ scenario is more problematic because many workers have no pension at all, while others have very generous defined benefit plans. To reflect this heterogeneity, the results are illustrated for three scenarios for  $y_R$ : no pension income, the highest pension income possible (\$4,000 for LHS, \$6,000 for HS, and \$24,000 for college), and a mid-point.<sup>17</sup>

The payroll tax rate  $\pi$  is 7.65%. The tax brackets  $[B_k, B_{k+1})$  and marginal tax rates  $\tau_k$  are taken from schedule X of the 2010 IRS 1040 form for singles, which is reproduced in Table 1. The standard deduction is \$5,700 and there is an additional \$1,400 deduction for singles above age 65. Since the personal exemption for a single person is \$3,650, a total of  $E = \$9,350$  and  $E_R = \$10,750$  is excludible from taxable income. The base amounts for the taxation of Social Security benefits are  $X^1 = \$25,000$  and  $X^2 = \$34,000$ , they vary over time according to (7). The tax brackets, deductions, and bend points in the Social Security benefits formula are assumed to increase with inflation in the future. The base amounts used in the taxation of Social Security benefits are fixed; an inflation rate of 3% is used to reflect their decrease in

A third-degree polynomial is fitted to the results. The resulting income profiles are similar to those reported in 2010 dollars in Brown *et al.* (2011).

<sup>16</sup> The PIA (Primary Insurance Amount) is 90% of the AIME (Average Indexed Monthly Earnings) up to \$761, plus 32% of the excess up to \$4,856, and 15% of the excess over \$4,856. The AIME is computed by taking the average of the highest 35 years of income indexed with average growth in wages. For simplicity, we assume that this is equal to inflation. The PIA is reduced by 6.66% because the individual retires at age 65 and it is assumed that the full retirement age is 66.

<sup>17</sup> The incentive to save for retirement can be lost if the combination of pension income with Social Security is such that replacement rates are too high. The ‘highest pension income possible’ was determined by finding the highest  $y_R$  such that the problem still had a solution with positive retirement savings.

Table 1. *Marginal tax rates*

$k$	$\tau_k$	$B_k$	$B_{k+1}$
0	0 %	$-\infty$	0
1	10 %	0	\$8,375
2	15 %	\$8,375	\$34,000
3	25 %	\$34,000	\$82,400
4	28 %	\$82,400	\$171,850
5	33 %	\$171,850	\$373,650
6	35 %	\$373,650	$\infty$

Source: Schedule X of the 2010 IRS 1040 form for singles.

real terms. To reflect these differences, the results are presented for individuals who were respectively aged 25, 45, and 65 in 2010.

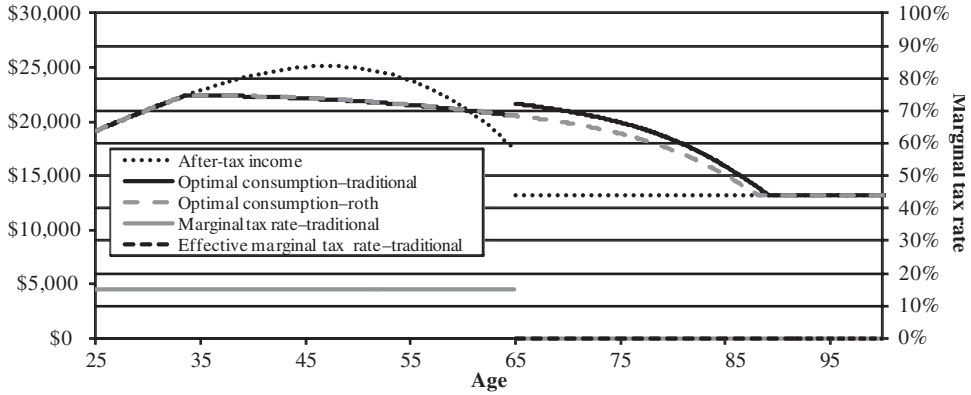
### 5 Endogenous tax rates and optimal consumption patterns

With standard problems, marginal tax rates are exogenously given. In contrast, the rates that apply to traditional accounts in this model are endogenously determined and can vary with time. To give the reader a sense of these rates, the bottom part of Figure 3 graphs the marginal tax rates  $\tau_k$  (straight lines) and the effective marginal tax rates  $\tau_k(1 + M_h)$  (dashed lines) for the traditional account solution as a function of age. Six representative cases with different education, cohort, and pension income are considered: Case 1 (no tax after retirement), Case 2 (few taxes after retirement), Case 3 (mixed rates after retirement), Case 4 (slightly higher taxes after retirement), Case 5 (much higher taxes after retirement), and Case 6 (slightly higher taxes after retirement).

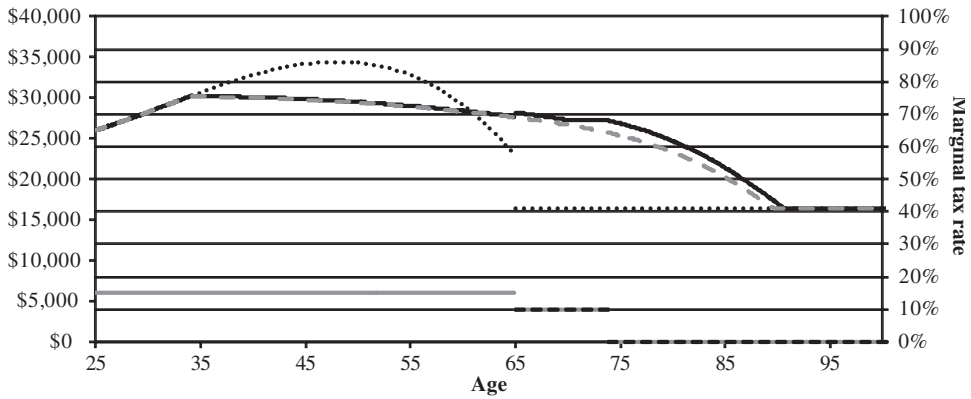
In our illustrations, the marginal tax rate before retirement is 15 % for the LHS/HS groups and 25 % for the college group. After retirement, effective marginal tax rates can be higher or lower and several rates can apply to the same person. The results in Figure 3 underscore the role played by the taxation of Social Security benefits: without it, Roth accounts would never be strictly preferred in our illustrations as the marginal tax rates  $\tau_k$  either stay the same or decrease after retirement. When  $M_h$  is taken into account, in many cases the resulting effective rates are above the pre-retirement rates, making Roth accounts potentially attractive. The question then is who is affected by the taxation of Social Security benefits. Currently, it is mostly those with higher incomes – in our illustrations, those in the college group. However, Social Security's base amounts are not indexed for inflation, which means that the situation will eventually apply to those in the high-school group for younger cohorts.

To show the connection between marginal tax rates and consumption, the top portion of Figure 3 gives the optimal consumption patterns for the traditional and Roth cases. In the figure, the link is evident: consumption is higher when marginal tax rates are lower (and vice-versa). Although the optimal consumption patterns for the traditional and Roth cases are similar in the sense that savings start and end around

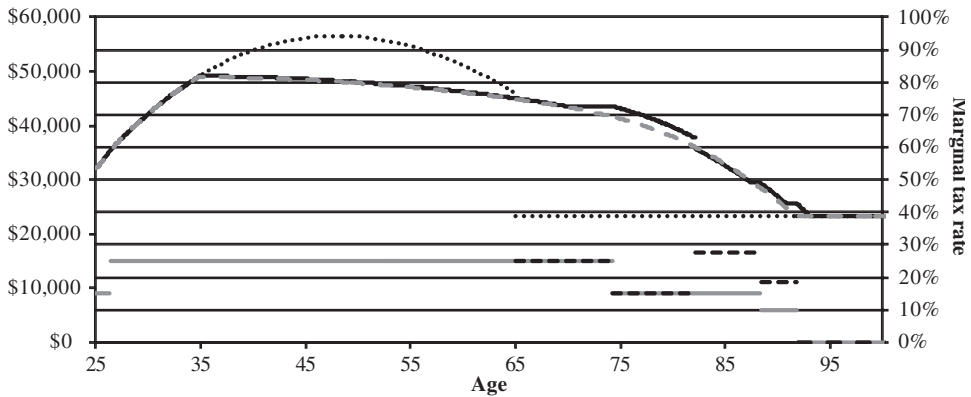
Case 1: No high school,  $y_R = \$0$ , age 45 in 2010



Case 2: High school,  $y_R = \$0$ , age 65 in 2010



Case 3: College,  $y_R = \$0$ , age 25 in 2010





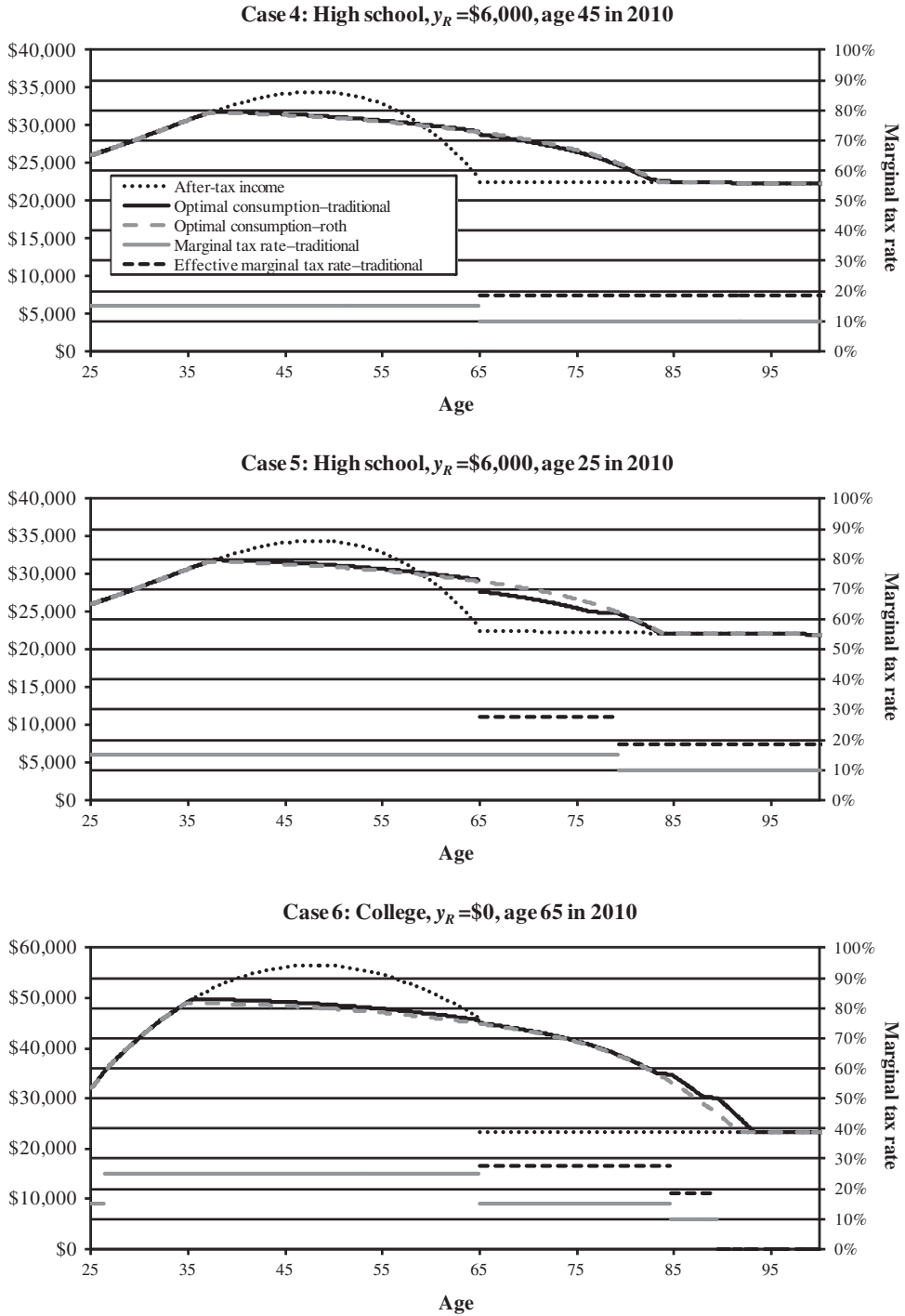


Figure 3. Examples of optimal consumption patterns.

the same ages, they differ in that the Roth solution is smooth and the traditional solution is more jagged.<sup>18</sup> As explained in Section 3, the solution goes through flat portions (at  $C_k$  or  $C_h$ ) when there is a transition between two tax rates, which can be observed in Cases 2, 3, 5, and 6. When  $M_h$  changes from 0% to 85%, it is also possible to have a jump in consumption: Case 3 provides an example of this when consumption suddenly declines by \$2,000 at age 82.

### 6 Who gains by choosing Roth accounts?

If an individual can only invest in a traditional or a Roth account, which one of the two is most beneficial? To answer that question, the first column of Table 2 presents a dollar measure of the welfare gains/losses that traditional accounts generate over their Roth counterparts. Appendix C details the equation used for that computation and Table 2 gives the results for each of the education groups (LHS, HS, and college), for three different cohorts (ages 25, 45, and 65 in 2010), and for three levels of pension income.<sup>19</sup>

The results in Table 2 indicate that in most cases choosing traditional accounts over Roths generates a welfare gain. The exceptions occur mainly when the taxation of Social Security benefits pushes the effective marginal tax rates after retirement above their pre-retirement levels, as illustrated in Cases 3–6 of Figure 3.<sup>20</sup> As discussed in the previous section, this mostly affects those in the college group for the current cohort of retirees, but will eventually affect those in the high-school group for the youngest generation. For example, Roth accounts are preferred by those who are currently 25 years old with  $y_R = \$6,000$ .<sup>21</sup> Those in the youngest generation who still prefer traditional accounts are also affected by the taxation of Social Security benefits as it cuts their welfare gains by about half.

#### 6.1 Breakeven increase in future tax rates

Since an increase in future tax rates would favor Roth accounts, it is interesting to ask: What is the magnitude of the increase needed to change the traditional account recommendation? To answer that question, recall that the model in Section 2 allows for a multiplicative increase  $\theta$  that applies to all marginal tax rates after retirement. Using this framework, we solve for the breakeven increase  $\theta$  that would make the individual indifferent between traditional and Roth accounts and present the results in the second column of Table 2. Note that the breakeven rates are negative for those who initially prefer Roth accounts. The table indicates ‘N/A’ for those who pay no (or very little) taxes after retirement since changes in future marginal tax rates are not

<sup>18</sup> For example, for someone age 45 in 2010 with no pension income, the respective values of  $T_1$  in the Roth and traditional cases are: 33.8 versus 33.7 (LHS), 33.9 versus 34.5 (HS), and 34.8 versus 35.2 (college). For  $T_2$ , the corresponding values are 88.1 versus 89.1 (LHS), 89.5 versus 90.6 (HS), and 91.8 versus 93.1 (college). For other cases, differences are typically within a year.

<sup>19</sup> It should be noted that in practice, those in the older cohorts did not have access to traditional and Roth accounts for their entire career.

<sup>20</sup> Traditional accounts can be optimal in some cases with higher marginal tax rates because different marginal tax rates can apply to withdrawals.

<sup>21</sup> At the same time, those with a college degree and  $y_R = \$24,000$  no longer prefer the Roth because they reached the maximum amount of taxable benefits and return to regular tax rates of 15% and 25%.

an issue for them, they will prefer traditional accounts no matter what. Future tax changes, however, are relevant for those with a college degree and in a few of the high-school cases. Table 2 shows that for them, breakeven rates range between 4% and 41%. Those with high pensions have low breakeven rates and can be affected by even small changes in tax rates. In contrast, those with a college degree and no pension income are less sensitive and would require a more substantial increase of 41% (age 45 in 2010) or 33% (age 25 in 2010) to justify the switch to Roth accounts. To put this in perspective, a 33% increase would change the marginal tax rates in Table 1 to 0%, 13%, 20%, 33%, 37%, 44%, and 47%.

Although Roth accounts are immune to increases in future marginal tax rates, it should be acknowledged that they are subject to different risks. For example, a change in the tax law could include withdrawals from Roth accounts in the taxation of Social Security benefits. To illustrate the magnitude of this issue, the third column of Table 2 gives the resulting increase in the present value of taxes for those who save with Roth accounts. This amount can be quite substantial, reaching up to \$18,443. If this change was expected with certainty, Roth accounts would never be optimal in our illustrations. In addition, Roth accounts would lose some of their appeal for younger cohorts if Social Security's base amounts become indexed for inflation. Of course, substantial tax reforms would also affect the comparison, for example Kotlikoff *et al.* (2008) evaluate the impact of a change to a consumption tax. Unfortunately, most strategies for retirement savings are not truly risk-free as long as the tax code can change and it is difficult to assign a probability distribution to these changes.

## 6.2 Naïve diversification strategies

Discussions of tax risk in the context of retirement savings are often accompanied by a suggestion to diversify among saving vehicles. For example, in a Vanguard document, Ahern *et al.* (2005) state: 'Pre-tax savings are more beneficial if a participant is in a lower tax bracket in retirement; Roth savings are more beneficial if a participant is in a higher bracket. In a world of uncertain future tax rates, participants should diversify. Just as they hold fixed income assets to diversify the risks of stocks, so participants should hold Roth savings to diversify the risks associated with pre-tax savings'. To investigate the potential benefits of having both types of accounts, we use the naïve diversification strategy defined in Section 2 where the individual allocates a constant proportion  $\alpha$  of savings/withdrawals to the traditional account and a proportion  $1 - \alpha$  to the Roth account. Of course, this strategy is limited in the sense that the individual cannot adjust  $\alpha$  in every period, but it gives us a simple platform to assess the welfare gains stemming from risk reduction. In the next subsection, we will discuss how mixed strategies can be improved over naïve ones.

Before evaluating the risk reduction benefits associated with a mixed strategy, we must first recognize that this approach can increase welfare even in the absence of risk. To understand why, consider the following example where the marginal tax rates before and after retirement are respectively 15% and 10%, but withdrawals in excess of \$10,000 trigger the taxation of Social Security benefits and are subject to an

Table 2. *Welfare gains/losses and related measures without tax risk*

Education/cohort/pension income after retirement	Pure strategies			Naïve diversification		
	Welfare gain/loss from traditional accounts (\$)	Breakeven increase in future tax rates <sup>1</sup> ( $\theta - 1$ )	Loss if Roth withdrawals included in $SS_t^{LX2}$ (\$)	Optimal fraction $\alpha$ in traditional accounts	Welfare gain/loss with optimal $\alpha$ (\$)	Value of option to invest in Roth accounts (\$)
<b>Less than high school</b>						
<b>Age 65 in 2010</b>						
$y_R = \$0$	14,184	N/A	0	100 %	14,184	0
$y_R = \$2,000$	10,250	N/A	0	100 %	10,250	0
$y_R = \$4,000$	6,677	N/A	0	100 %	6,677	0
<b>Age 45 in 2010</b>						
$y_R = \$0$	14,184	N/A	0	100 %	14,184	0
$y_R = \$2,000$	10,250	N/A	0	100 %	10,250	0
$y_R = \$4,000$	6,677	N/A	0	100 %	6,677	0
<b>Age 25 in 2010</b>						
$y_R = \$0$	12,669	N/A	0	100 %	12,669	0
$y_R = \$2,000$	7,873	N/A	0	90 %	7,899	26
$y_R = \$4,000$	3,387	N/A	0	70 %	3,675	288
<b>High school</b>						
<b>Age 65 in 2010</b>						
$y_R = \$0$	22,020	N/A	0	100 %	22,020	0
$y_R = \$3,000$	15,445	N/A	0	100 %	15,445	0
$y_R = \$6,000$	9,300	N/A	0	100 %	9,300	0
<b>Age 45 in 2010</b>						
$y_R = \$0$	17,114	N/A	0	90 %	17,302	188
$y_R = \$3,000$	9,357	N/A	0	70 %	10,034	677
$y_R = \$6,000$	2,132	28 %	799	40 %	3,465	1,333

Roth versus traditional accounts in a life-cycle model with tax risk

**Age 25 in 2010**

$y_R = \$0$	9,616	N/A	0	60%	11,384	1,768
$y_R = \$3,000$	1,042	7%	2,439	30%	4,486	3,444
$y_R = \$6,000$	-4,518	-29%	5,418	0%	0	4,518

**College****Age 65 in 2010**

$y_R = \$0$	47,556	N/A	0	90%	48,128	572
$y_R = \$12,000$	1,601	4%	10,819	40%	4,074	2,473
$y_R = \$24,000$	-1,418	-10%	5,120	0%	0	1,418

**Age 45 in 2010**

$y_R = \$0$	26,461	41%	5,806	60%	29,383	2,922
$y_R = \$12,000$	-4,388	-8%	18,443	0%	0	4,388
$y_R = \$24,000$	-393	-2%	1,551	20%	126	519

**Age 25 in 2010**

$y_R = \$0$	22,694	33%	12,761	100%	22,694	0
$y_R = \$12,000$	5,308	12%	11,988	100%	5,308	0
$y_R = \$24,000$	1,160	7%	0	100%	1,160	0

<sup>1</sup> This is the percent increase in marginal tax rates after retirement which would make the individual indifferent between the traditional and Roth accounts. N/A indicates that the individual does not pay taxes after retirement or so little that it would take an increase over 80% to make the individual indifferent.

<sup>2</sup> For the case where the individual invests all his savings in a Roth account, this column gives the present value of the increase in taxes that would result if withdrawals from Roth accounts would count in the taxation of Social Security benefits.

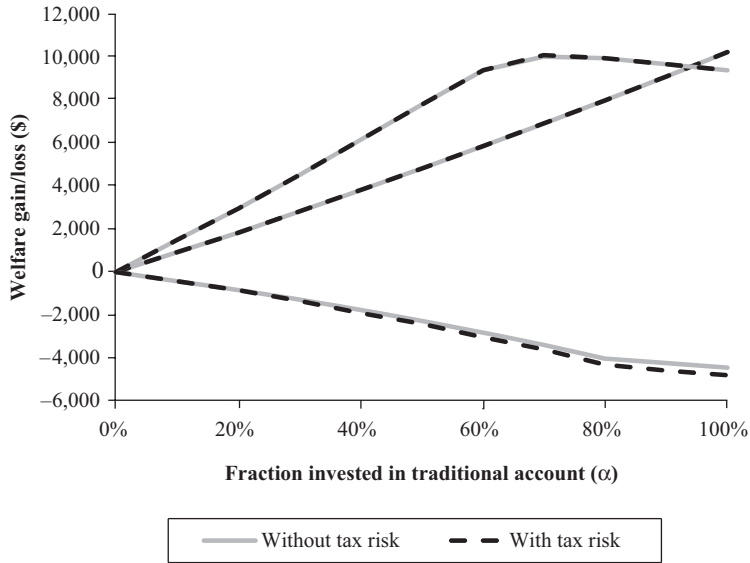


Figure 4. Welfare gains with naïve diversification strategies (over Roth accounts) for selected cases (age 45 in 2010).

effective marginal tax rate of 18.5%. An individual who has to make annual withdrawals of \$15,000 can use a naïve strategy with  $\alpha = 2/3$  and withdraw \$10,000 from a traditional account and \$5,000 from a Roth account. This strategy is beneficial because it allows the individual to gain from the lower tax rate of 10%, while avoiding the higher rate of 18.5%.

For the case without tax risk, we compute the welfare gains associated with each  $\alpha$  from 0 to 1 and thus are able to find the optimal  $\alpha$ . The results are presented in the fourth and fifth columns of Table 2 for each of the cases considered previously. Figure 4 also gives a graphical representation of the welfare gains as a function of  $\alpha$  for three representative cases. Those who do not encounter issues with the taxation of Social Security find it optimal to allocate 100% of savings to traditional accounts. A 100% Roth strategy is optimal for those whose effective marginal tax rates after retirement are higher than before retirement even before making any withdrawal. For those who are in the situation described in the previous paragraph (about 40% of our cases), it is optimal to divert some (but not all) savings to the Roth account to avoid the higher marginal tax rates. Note that the welfare gains with the optimal  $\alpha$  are always non-negative: losses with traditional accounts can be avoided when there is an option to allocate part of savings to a Roth account. The value of the option to invest a fraction of savings in Roth accounts can be computed by taking the difference between the two welfare gains in Table 2: the last column of Table 2 shows that this value is on average \$1,800 when positive, with a maximum of \$4,518.<sup>22</sup>

<sup>22</sup> These figures are a lower bound for the value of the option given that we are using a naïve diversification strategy instead of optimizing  $\alpha$  in every period.

We now introduce tax variability with a simple no-drift scenario where marginal tax rates after retirement can go up or down by 20%. Using Section 2's notation, this translates into  $p_1 = p_2 = 50\%$ ,  $\theta_1 = 120\%$ , and  $\theta_2 = 80\%$ . Figure 4 contrasts the welfare gains for the cases with risk (dashed lines) and without tax risk (solid lines). For those in the LHS/HS education groups, the lines coincide and tax risk has essentially no impact on welfare gains. For those in the college group, we observe a small difference and tax risk reduces welfare by at most \$409 when  $\alpha = 100\%$ . The loss attributable to tax risk diminishes with diversification, for example it is cut by \$281 when  $\alpha = 50\%$ . However, Figure 4 shows that risk reduction benefits are not the only consideration when choosing  $\alpha$ : gains or losses in the scenario without risk must also be taken into account. In the previous example with  $\alpha = 50\%$ , diversification is suboptimal because the certain loss (\$1,580) is much higher than the risk reduction benefits (\$281).

More generally, in most of our illustrations the magnitude of the risk reduction benefit that a given  $\alpha$  brings is much smaller than the corresponding gain or loss in the no-risk case. Indeed, the optimal proportions allocated to traditional accounts in Table 2 are essentially the same for the cases with and without tax risk. Although intuitive at first, the analogy with a diversified portfolio of stocks and bonds does not translate well because traditional and Roth accounts can have great differences in expected values, which dominate the volatility effect. On the other hand, the certainty case showed that the peculiar nature of the tax structure provides a new motivation for diversifying: a mix of traditional and Roth accounts can have a higher expected value than a linear combination of the pure cases.

### 6.3 Mixed strategies

The naïve diversification strategy suggested in the previous section can be fine-tuned to further improve welfare. For instance, instead of mixing both accounts in every period before retirement, the optimal strategy is likely to involve a switch between periods where contributions are either 100% traditional or 100% Roth.<sup>23</sup> Unless there are major fluctuations in income, there should be relatively few transitions between the two accounts. Starting with the Roth account can be justified by lower marginal tax rates at that time or by concerns about early withdrawals and the 10% penalty tax.<sup>24</sup> Conversely, starting with the traditional account is preferable when these are not an issue and marginal tax rates decline after retirement.<sup>25</sup> If savings in the traditional account eventually reach a level such that the marginal tax rate after retirement exceeds the pre-retirement one, a permanent switch to Roth accounts would be recommended. Temporary moves to the Roth side could also be motivated in periods of income loss or with particularly high tax deductions. After retirement, the strategy would be a generalization of our earlier example with  $\alpha = 2/3$  to every

<sup>23</sup> Issues with maximum contributions or eligibility could make partial strategies optimal.

<sup>24</sup> Although the 10% penalty tax applies to the entire amount of the withdrawal for a traditional account, for Roth accounts only the investment income portion is subject to the penalty.

<sup>25</sup> To see this, let  $\tau_1$  and  $\tau_2$  be the marginal tax rates at times  $t_1$  and  $t_2$ . A dollar invested at time  $t_1$  and withdrawn at time  $t_2$  yields  $e^{r(t_2-t_1)}(1-\tau_2)/(1-\tau_1)$  in the traditional case and  $e^{r(t_2-t_1)}$  in the Roth case. If  $\tau_2 < \tau_1$ , it is advantageous to save in the traditional account for a longer period of time  $t_2 - t_1$ .

period: withdrawals would be made from the traditional account first, and  $\alpha$  would be chosen such that higher marginal tax rates are avoided.

**7 Retirement savings**

The next question investigated is whether tax deductible contributions increase retirement savings. If so, is it an income or a substitution effect? In other words, savings may increase either because they are augmented by a tax subsidy or because people sacrifice more consumption before retirement. For this discussion, a measure of gross retirement savings can be obtained by computing the accumulated value of savings as follows:

$$\text{Gross retirement savings} = W_R = \int_{T_1}^R e^{r(R-t)} s(c_t^*) dt. \tag{23}$$

In the traditional account case, gross retirement savings are inflated in the sense that taxes will have to be paid on withdrawals. Accordingly, we also define a net measure of retirement savings where a ‘tax liability’ is deducted. This tax liability measures the present value of taxes attributable to withdrawals ( $s_t$ ) and it can be obtained by taking the difference between taxes with and without withdrawals as follows:

$$\begin{aligned} \text{Tax liability} &= \int_R^\omega e^{-r(t-R)} [\text{tax}_t(s_t) - \text{tax}_t(0)] dt \\ \text{Net retirement savings} &= W_R - \text{Tax liability}. \end{aligned} \tag{24}$$

**7.1 Increases in retirement savings**

Table 3 presents the gross and net levels of retirement savings at age 65 for each scenario, assuming that the individual was aged 45 in 2010. By education group, these range from: \$40,000–\$100,000 (LHS), \$60,000–\$140,000 (HS), and \$50,000–\$350,000 (college). Not surprisingly, higher income translates into higher savings, whereas higher pensions reduce the need to accumulate wealth. The bottom part of Table 3 tests the sensitivity of these results to the following changes in parameters:  $a = 25$ ,  $a = 45$ ,  $r = 1\%$ ,  $r = 5\%$ ,  $\beta = 1$ ,  $\beta = 5$ ,  $\gamma = 1$ ,  $\gamma = 5$ , a reduction of 25% in Social Security benefits, and an increase by \$5,000 in the exemptions amounts  $E$  and  $E_R$ . All cases are considered in the analysis, but due to space limitations the tables present only the averages for all education/income categories. The results show that the level of retirement savings is very sensitive to the choice of parameters, being halved or almost doubled in some scenarios.

The more salient result in Table 3 is the marginal impact of traditional accounts on retirement savings. First, those in the less-than-high-school category increase their savings by about \$15,000. In their case, the gross and net differences are the same because they do not have enough income after retirement to pay taxes. The differences for the high-school group are much smaller: the gross increase is about \$6,000 and the net increase averages only \$200. Finally, the college group displays the more striking results. In their case, the gross increase can be pretty substantial reaching up



Table 3. *Retirement savings at age 65 (in dollars; age 45 in 2010)*

Education/ pension income after retirement	Retirement savings		Tax liability for traditional account	Increase in retirement savings with traditional accounts	
	Roth	Traditional		Gross	Net of tax liability
<b>Less than high school</b>					
$y_R = \$0$	84,958	101,717	0	16,759	16,759
$y_R = \$2,000$	59,859	75,197	0	15,338	15,338
$y_R = \$4,000$	37,249	50,933	0	13,684	13,684
<b>High school</b>					
$y_R = \$0$	136,113	143,973	4,058	7,860	3,802
$y_R = \$3,000$	97,145	103,281	5,853	6,136	283
$y_R = \$6,000$	61,997	66,206	7,713	4,209	-3,503
<b>College</b>					
$y_R = \$0$	284,473	352,400	61,027	67,927	6,899
$y_R = \$12,000$	144,744	199,593	54,059	54,849	790
$y_R = \$24,000$	51,898	70,590	17,982	18,692	710
Sensitivity analysis (averages for all education/pension income categories)					
Baseline	106,493	129,320	16,744	22,828	6,085
$a = 25$	107,739	126,794	19,016	19,055	39
$a = 65$	104,068	124,037	10,028	19,969	9,943
$r = 1\%$	63,123	77,483	7,074	14,360	7,288
$r = 5\%$	172,099	200,111	33,296	28,012	-5,284
$\beta = 1$	183,065	212,856	33,259	29,791	-3,467
$\beta = 5$	61,946	74,765	7,459	12,819	5,361
$\gamma = 1$	44,351	60,901	4,359	16,550	12,191
$\gamma = 5$	133,008	160,010	22,009	27,002	4,993
SS cut by 25%	161,647	188,542	21,912	26,896	4,984
$E = \$14,350$	115,830	140,261	18,536	24,432	5,896
$E_R = \$15,750$	104,922	129,361	11,857	24,439	12,582

to \$68,000. However, once the increase in the tax liability is considered, this gain mostly vanishes with an average of only \$2,800. Testing the sensitivity of these results, we find that the meaningful increase in net retirement savings for the less-than-high-school group is generally robust. The high-school group exhibits more variable results, ranging from -\$20,000 to \$20,000. The college group displays a pattern mostly similar to the one observed in the top portion of Table 3: a high increase in gross retirement savings is experienced, but most of it goes away once the net values are considered.

## 7.2 *Income and substitution effects*

Traditional accounts are more effective at increasing savings in some cases than others. To understand why, this section suggests a breakdown for the increase in

retirement consumption (or equivalently the increase in net retirement savings) into an income effect and a substitution effect. The income effect comes from the additional tax subsidy that tax deductible contributions generate. It can be computed as the difference in the lifetime value of taxes as follows:

$$TS = \text{Tax subsidy} = \int_{t_0}^{\omega} e^{-r(R-t)} [\text{tax}_t^{\text{Roth}}(s_t) - \text{tax}_t^{\text{Trad}}(s_t)] dt. \tag{25}$$

It should be noted that retirement consumption does not increase by the entire amount of the tax subsidy, the wealth effect increases consumption in all periods, both before and after retirement.<sup>26</sup> Introducing the notation  $C_{t,T} = \int_t^T e^{-r(R-s)} c_s^* ds$ , the proportion of the tax subsidy allocated to retirement consumption can be estimated with the following ratio:

$$q = \text{Propensity to consume after retirement} = \frac{C_{R,\omega}}{C_{t_0,\omega}}. \tag{26}$$

A reduction in tax rates after retirement not only creates a subsidy but also a substitution effect by lowering the relative price of post-retirement consumption. The value of the new savings before retirement associated with the substitution effect can be computed with  $\text{New Savings} = C_{t_0,\omega}^{\text{Roth}}(q^{\text{Trad}} - q^{\text{Roth}})$ . Accordingly, the total changes in pre- and post-retirement consumption<sup>27</sup> can be written as

$$\begin{aligned} \text{Change in pre-retirement consumption} &= (1 - q^{\text{Trad}}) \cdot TS - \text{New Savings} \\ \text{Change in retirement consumption} &= q^{\text{Trad}} \cdot TS + \text{New Savings} \end{aligned} \tag{27}$$

Table 4 presents the results of the decomposition for the change in retirement consumption along with the tax subsidies and the applicable effective marginal tax rates. The table includes the previous cases from Figure 3 and we will use them as representative examples. The average tax subsidy in Table 4 is about \$10,000; since  $q^{\text{Trad}}$  is generally around 20 %, the average increase in retirement consumption attributable to the tax subsidy is only \$2,000. By themselves, subsidies can be a relatively expensive way to generate a small increase in retirement consumption. For example, the increase is at best \$9,240 in Case 6, but the total cost of the subsidy is \$48,396.

The substitution effect can generate larger increases in retirement consumption than the income effect, but only for some groups. It works well for those who pay little or no taxes after retirement, for example in Cases 1 and 2 the associated increases in retirement consumption are \$13,745 and \$14,248. For those who pay meaningful taxes after retirement, the taxation of Social Security benefits again changes the cards and reduces the substitution effect’s potential. In many cases, marginal tax rates increase and the substitution effect is actually negative.

<sup>26</sup> In the solution presented in Section 3, an increase in wealth reduces  $\lambda$  and consequently increases consumption in all periods.

<sup>27</sup> Note that this decomposition indicates that the impact of a policy on saving behavior cannot be asserted simply by examining the change in pre-retirement consumption as this amount also includes a tax subsidy portion. For example, Gomes *et al.* (2009) concluded that tax-deferred accounts do not promote a greater propensity to save because they increased pre-retirement consumption by 2 % on average. If the 2 % includes a tax subsidy, a positive substitution effect may be present but hidden.

Table 4. Increase in retirement consumption

Education/ cohort/pension income after retirement	Increase in retirement consumption			Effective marginal tax rates after retirement $\tau_k(1 + M_h)$	
	Tax subsidy	From tax subsidy (income)	From new savings (substitution)		Total
<b>Less than high school</b>					
<b>Age 65 in 2010</b>					
$y_R = \$0$	15,258	3,014	13,745	16,759	0%
$y_R = \$2,000$	11,280	2,290	13,047	15,338	0%
$y_R = \$4,000$	7,640	1,599	12,084	13,684	0%
<b>Age 45 in 2010</b>					
$y_R = \$0$	15,258	3,014	13,745	16,759	0%
$y_R = \$2,000$	11,280	2,290	13,047	15,338	0%
$y_R = \$4,000$	7,640	1,599	12,084	13,684	0%
<b>Age 25 in 2010</b>					
$y_R = \$0$	13,117	2,522	2,900	5,422	18.5%, 0%
$y_R = \$2,000$	8,200	1,617	824	2,442	18.5%, 0%
$y_R = \$4,000$	3,562	724	-1,234	-510	18.5%, 0%
<b>High school</b>					
<b>Age 65 in 2010</b>					
$y_R = \$0$	22,984	4,451	14,248	18,698	10%, 0%
$y_R = \$3,000$	16,177	3,217	11,644	14,861	10%, 15%, 0%
$y_R = \$6,000$	9,769	2,002	8,502	10,505	10%, 15%, 0%
<b>Age 45 in 2010</b>					
$y_R = \$0$	17,538	3,309	493	3,802	18.5%, 0%
$y_R = \$3,000$	9,639	1,872	-1,589	283	18.5%, 0%
$y_R = \$6,000$	2,217	444	-3,949	-3,503	18.5%
<b>Age 25 in 2010</b>					
$y_R = \$0$	10,560	1,961	-7,625	-5,644	27.75%, 18.5%, 0%
$y_R = \$3,000$	2,057	392	-11,353	-10,961	27.75%, 18.5%, 0%
$y_R = \$6,000$	-3,464	-680	-13,241	-13,921	27.75%, 18.5%
<b>College</b>					
<b>Age 65 in 2010</b>					
$y_R = \$0$	48,396	9,240	1,664	10,904	27.75%, 18.5%, 0%
$y_R = \$12,000$	1,702	346	-6,206	-5,857	27.75%
$y_R = \$24,000$	-1,323	-289	-5,115	-5,404	27.75%
<b>Age 45 in 2010</b>					
$y_R = \$0$	27,073	5,170	1,730	6,899	46.25%, 27.75%, 25%, 18.5%, 15%, 0%
$y_R = \$12,000$	-4,161	-848	1,637	790	25%, 15%, 27.75%
$y_R = \$24,000$	-355	-73	783	710	25%
<b>Age 25 in 2010</b>					
$y_R = \$0$	23,521	4,541	10,842	15,382	25%, 15%, 27.75%, 18.5%, 0%
$y_R = \$12,000$	5,747	1,173	5,930	7,103	25%, 15%, 27.75%
$y_R = \$24,000$	1,218	267	794	1,061	25%

For example, in Case 5 the marginal tax rate after retirement jumps to 27.75% and retirement consumption is reduced by \$13,214.

To conclude this section, it is interesting to put the tax subsidies in perspective by observing that they move in sync with the welfare gains in Table 2. In other words, welfare gains associated with the tax deductibility of contributions arise because they are paid for by tax subsidies. Tax deductible contributions do not generate a benefit above their cost: on average, welfare gains are \$500 less than tax subsidies because consumption patterns are disrupted. Although tax deductible contributions benefit most people in partial equilibrium, it should be kept in mind that this is not necessarily the case in general equilibrium where tax subsidies have to be financed by an increase in other taxes.

## 8 Conclusion

The results in this paper have a number of interesting applications for financial planners and 401(k)s. For financial planners who advise clients on their Roth/traditional decision, it is a good news, bad news story. The bad news is that the case for which we can determine with more conviction that traditional accounts are superior is for those who have lower incomes (i.e. those who are less likely to seek advice.) For those with higher incomes and some pensions, the results are often not as clear cut – the good news is that welfare losses from making the wrong decision are not excessively high. Actually, we show that a mixed traditional/Roth strategy can improve welfare when it helps to avoid the higher marginal tax rates due to the taxation of Social Security benefits. In contrast, we find that naïve diversification strategies offer limited risk reduction benefits when tax risk takes only the form of variability in future tax rates.

Our findings have also applications in the realm of 401(k)s: recently, employees were given the possibility of directing their 401(k) contributions to a Roth account, but employer matching contributions can only go to a traditional account. As some employees can lose with traditional accounts, our results suggest that extending the Roth opportunity to employer contributions would benefit these employees. Moreover, default strategies are becoming increasingly important in 401(k) plans. Although these have focused on contribution levels and asset allocation, it would also be interesting to consider which type of retirement account (traditional or Roth) should be set as a default option. This paper illustrated results for a wide range of cases and offers a starting point for this type of analysis. In particular, this paper hopes to raise awareness in terms of the underappreciated role played by the taxation of Social Security benefits.

This paper also offers new developments from a methodological perspective. By exploiting a dual approach, we illustrate how an analytical framework can be retained even when incorporating elements such as borrowing constraints, risks, and discontinuities. The analytical approach also provides valuable new insight by showing that the solution can be based on a system of known budget constraint equations. The benefit of this alternative formulation over more conventional dynamic programming is twofold: it produces exact values and eliminates the

time-consuming process of estimating an unknown value function with numerical optimization and backward induction.

Finally, this paper found that the differences between Roth and traditional accounts are limited to some groups within the context of a standard life-cycle model analysis. The interesting question then is whether there are also practical considerations that affect the comparison. For example, the tax deductibility of contributions may be associated with behavioral effects that lead people to save more than they would with Roth accounts. That could be the case if the value of the immediate tax refund looms larger than the associated tax liability in the decision. Another intriguing issue is the fact that the tax deductibility of contribution can increase gross savings significantly. This is advantageous for an individual who is able to achieve superior investment returns. Similarly, having more assets under management is obviously beneficial for the industry. On the other hand, postponing tax receipts might not be desirable for cash-strapped governments. In addition, unsophisticated investors are prone to investment mistakes and this problem is leveraged with larger assets. These potential issues are left for future research.

### References

- Ahern, M., Americks, J., Dickson, J., Nestor, R. and Utkus, S. (2005) Tax diversification and the Roth 401(k). *Vanguard Center for Retirement Research*, 18.
- Boyd, S. and Vandenberghe, L. (2004) *Convex Optimization*. Cambridge: Cambridge University Press.
- Brown, J.R., Fang, C. and Gomes, F. (2011) Risk and returns to education: The value of human capital. Working Paper, University of Illinois at Urbana-Champaign.
- Burman, L.E., Gale, W.G. and Weiner, D. (2001) The taxation of retirement saving: choosing between front-loaded and back-loaded options. *National Tax Journal*, **54**(3): 689–702.
- Butterfield, S.L., Jacobs, F.A. and Larkins, E.R. (2000) The Roth versus the traditional IRA: a comparative analysis. *Journal of Applied Business Research*, **16**(4): 113–128.
- Cocco, J., Gomes, F. and Maenhout, P. (2005) Consumption and portfolio choice over the life cycle. *Review of Financial Studies*, **18**(2): 491–533.
- Dammon, R.M. (2009) Retirement investing: analyzing the Roth conversion option. Tepper School of Business, Paper 104.
- Dickson, J.M. (2004) Pension taxation and tax code risk. In E.G. Gale, J.B. Shoven, and M.J. Warshawsky (eds). *Private Pensions and Public Policies*, Brookings Institution Press, 221–231.
- Engen, E.M., Gale, W.G. and Scholz, J.K. (1994) Do savings incentive work? *Brookings Papers on Economic Activity*, **25**(1): 85–180.
- Gokhale, J., Kotlikoff, L.J. and Neumann, T. (2001) Does participating in a 401(k) raise your lifetime taxes? National Bureau of Economic Research, Working paper 8341.
- Gomes, F., Michaelides, A. and Polkovnichenko, V. (2009) Optimal savings with taxable and tax-deferred accounts. *Review of Economic Dynamics*, **12**: 718–735.
- He, H. and Pages, H.F. (1993) Labor income, borrowing constraints, and equilibrium asset prices. *Economic Theory*, **3**: 663–696.
- Kotlikoff, L.J., Marx, B. and Rapson, D. (2008) To Roth or not? That is the question. National Bureau of Economic Research, Working paper 13763.
- Lachance, M.E. (2012) Optimal onset and exhaustion of retirement savings in a life-cycle model. *Journal of Pension Economics and Finance*, **11**(1): 21–52.

Laibson, D.I., Repetto, A. and Tobachman, J. (1998) Self-control and saving for retirement. *Brookings Papers on Economic Activity*, 91–172.  
 Love, D.A. (2007) What can the life-cycle model tell us about 401(k) contributions and participation? *Journal of Pension Economics and Finance*, 6(2): 1473–1492.  
 Yaari, M.E. (1965) Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies*, 32(3): 137–150.

**Appendix A: Technical details for the solution of Section 2.3’s optimization problem**

*A.1 Standard Lagrangian and dual approach*

To prove that a solution to an optimization problem with constraints is optimal, the standard approach is to start by constructing a Lagrangian function where each constraint is multiplied by its Lagrange multiplier and the result appended to the objective function as follows:

$$L = E \left[ \int_{t_0}^{\omega} f(t)u(c_t)dt + \mu W_{t_0} + \int_{t_0}^{\omega} \eta_t W_t dt \right]. \tag{A.1}$$

In (A.1),  $W_t$  is the wealth process and  $\mu$  and  $\eta_t$  denote, respectively, the Lagrange multipliers for the budget constraint and the borrowing constraint at time  $t$ . In state  $i$  after retirement,  $p_i\mu_i$  and  $p_i\eta_t^i$  are used instead. To reduce clutter, we omit the subscripts  $i$  in this Appendix unless necessary. With this notation defined, an optimal solution must satisfy the following four Karush–Kuhn–Tucker (KKT) necessary conditions for all  $t$ : (1)  $W_0^* = 0$  and  $W_t^* \geq 0$ , (2)  $\eta_t \geq 0$ , (3)  $\eta_t W_t^* = 0$ , and (4)  $L'(c_t^*) = 0$ . As mentioned in Section 3, the solution  $c_t^*$  to  $L'(c_t^*) = 0$  gives us little insight in terms of showing that  $W_t^* \geq 0$  everywhere.

To overcome this problem, the dual approach suggested in Lachance (2012) is used and adapted to handle tax risk and discontinuities. Following He and Pages (1993) and applying integration by parts,<sup>28</sup> the Lagrangian in (A.1) can be rewritten in terms of a process  $X(t)$  as follows:

$$L = E \left[ \int_{t_0}^{\omega} (f(t)u(c_t) + X(t)e^{-r(t-t_0)}s_t(c_t)) dt \right],$$

where

$$X(t) = E \left[ \mu + \int_t^{\omega} \eta_s ds \right] \text{ for } t < R \quad \text{and} \quad X^i(t) = \mu_i + \int_t^{\omega} \eta_s^i ds \text{ for } t \geq R. \tag{A.2}$$

With this form, the dual approach can be applied as a two-step process: (1) find  $c_t^*$  that unconditionally maximizes the Lagrangian and (2) substitute the result in  $L$  to find the process  $X(t)$  that minimizes  $L$ .<sup>29</sup> The advantage of this formulation is that it does not require that we show that  $W_t^* \geq 0$  everywhere. The first step is easy because

<sup>28</sup> To write (A.2) in a form similar to He and Pages (1993), we substitute  $\int_{t_0}^t e^{r(t-t_0)}s_t dt$  to  $W_t$  in (28).  
<sup>29</sup> Under certain technical conditions (see Boyd and Vandenberghe, 2004), the value function with the dual approach is given by  $V^* = \inf_{X \in D} \sup_{c > 0} L(c) = \inf_{X \in D} L(c^*)$  where D is the set of non-negative and decreasing functions.

we solve an unconstrained problem instead of a constrained one; the solution for  $c_t^*$  is given in Section 3.2 and is not repeated here. The next section explains how to solve for  $X(t)$  in the second step.

### A.2 Solution for $X(t)$

*Binding periods:* In periods where the borrowing constraint is binding, the individual consumes as much as he can and  $c_t^* = \bar{y}_t = y_t - \text{tax}_t(0)$ . Substituting this in equation (17) and replacing  $\lambda$  by  $X(t)$ , we can invert the only possible solution  $X(t) = \lambda(t)$  where

$$\lambda(t) = u'(\bar{y}_t) f(t) e^{r(t-t_0)} (1 - \alpha \tau_k (1 + M_h)). \quad (\text{A.3})$$

Note that binding periods can only happen in periods when  $\lambda(t)$  is decreasing since  $X(t)$  must be decreasing in these periods.

*Non-binding periods:* Within a period where  $W_t^* > 0$ , by construction  $X(t)$  must be constant and we denote this by  $X(t) = \lambda$ .<sup>30</sup> With tax risk, a different constant  $\lambda_i$  is used in each state  $i$  after retirement. If there is more than one non-binding period, a different (and decreasing) constant  $\lambda$  would be used in each separate period. For simplicity, we use a single period  $[T_1, T_2]$  below, but the same concepts would apply with multiple periods.

*Connection points between periods:* At the connection points  $T_1$  and  $T_2$ , the solution is generally continuous and  $\lambda(T_1) = \lambda = \lambda(T_2)$ .<sup>31</sup> With risk, this condition becomes:

$$\lambda(T_1) = \lambda = \sum_{i=1}^N p_i \lambda_i = E[\lambda(T_2)] \quad \text{where} \quad \lambda_t = \lambda(T_2^t). \quad (\text{A.4})$$

Within intervals where  $\lambda(t)$  is strictly decreasing, this condition can be used to express  $T_1$  and  $T_2$  as inverse functions  $T_1(\lambda)$  and  $T_2(\lambda)$ . For the special case where  $T_1 = t_0$ , the condition  $\lambda(T_1) = \lambda$  becomes  $\lambda \geq \lambda(t_0)$ .

*Budget constraint and  $X(t)$  that minimizes  $L$ :* For each state  $i = 1, \dots, N$ , the budget constraint can be written as the present value of savings over the interval  $[T_1, T_2^i]$  with

$$W(\lambda, \lambda_i) = \int_{T_1(\lambda)}^R e^{-r(t-T_1(\lambda))} s(c_t^*(\lambda)) dt + \int_R^{T_2^i(\lambda_i)} e^{-r(t-T_1(\lambda))} s(c_t^*(\lambda_i)) dt = 0. \quad (\text{A.5})$$

Restricting the processes  $X(t)$  to those that satisfy the budget constraint, the second component in  $L$  is zero.  $L$  becomes:

$$L = E \left[ \int_{t_0}^{T_1(\lambda)} f(t) u(\bar{y}_t) dt + \int_{T_1(\lambda)}^{T_2(\lambda)} f(t) u(c_t^*(\lambda)) dt + \int_{T_2(\lambda)}^{\omega} f(t) u(\bar{y}_t) dt \right] \quad (\text{A.6})$$

and we can show that  $dL/d\lambda < 0$ . Since  $L$  decreases with  $\lambda$ , the criteria '  $X(t)$  that minimizes  $L$  ' reduces to choosing the process with the highest  $\lambda$  among those that satisfy the budget constraint. The next section gives a practical algorithm to solve for  $\lambda$ .

<sup>30</sup> This follows from the third KKT condition in appendix A.1 that requires that  $\eta_t = 0$  when  $W_t^* > 0$  and from the equation for  $X(t)$  in (A.2).

<sup>31</sup> This result follows from two conditions: (i) since  $X(t)$  is decreasing,  $\lambda(T_1^-) \geq \lambda$  and  $\lambda \geq \lambda(T_2)$ , and (ii) to ensure that the individual is saving (dissaving) at the beginning (end) of  $[T_1, T_2]$ , we must have  $\lambda > \lambda(T_1^+)$  and  $\lambda < \lambda(T_2^-)$ . If  $\lambda(t)$  is continuous at  $T_1$  and  $T_2$ , these conditions imply  $\lambda(T_1) = \lambda = \lambda(T_2)$ . If there is a discontinuity at  $T_1$  or  $T_2$ , the condition is adjusted to  $\lambda(T_1^-) \geq \lambda > \lambda(T_1^+)$  or  $\lambda(T_2^-) \geq \lambda > \lambda(T_2)$ .

### A.3 Algorithm to solve for $\lambda$

To develop a practical algorithm to solve for  $\lambda$ , the key is to formulate the problem as an equation  $g(\lambda)=0$  to which the bisection method can be applied. Recall that for an interval  $\lambda \in (\lambda_1, \lambda_2)$ , the intermediate value theorem guarantees the existence of a unique solution to  $g(\lambda)=0$  if  $g'(\lambda)>0$  for all  $\lambda \in (\lambda_1, \lambda_2)$ ,  $g(\lambda_1)<0$ , and  $g(\lambda_2)>0$ . From the previous section, our problem is to find values  $\lambda$  and  $\lambda_i$  such that the  $N$  budget constraint equations  $W(\lambda, \lambda_i)=0$  are satisfied, the condition  $\lambda = \sum_{i=1}^N p_i \lambda_i$  is met,  $\lambda \geq \lambda(t_0)$  if  $T_1 = t_0$ ,  $\lambda = \lambda(T_1)$  if  $T_1 > t_0$ ,  $\lambda_i = \lambda(T_2^i)$ , and  $\lambda'(t) < 0$  for all  $t \in [t_0, T_1) \cup [T_2^i, \omega]$ . For a given  $\lambda$ , retirement savings can be expressed as a function  $W_R(\lambda)$ . When  $W_R(\lambda) > 0$ , the budget constraint equation  $W(\lambda, \lambda_i) = 0$  allows us to express  $\lambda_i$  as an inverse function  $\lambda_i(\lambda)$  with  $\lambda_i'(\lambda) < 0$ .<sup>32</sup> If there is only one period  $[t_0, \tilde{t})$  before retirement where  $\lambda'(t) < 0$ , we can define a value  $\bar{\lambda} \in (\lambda(\tilde{t}), \infty)$  such that retirement wealth is positive for all  $\lambda > \bar{\lambda}$ .<sup>33</sup> This also allows us to define an inverse function  $T_1(\lambda)$  with  $\lambda'(t) < 0$  for all  $t \in [t_0, T_1)$ .

Our problem can then be formulated as a single equation  $g(\lambda) = \lambda - \sum_{i=1}^N p_i \lambda_i(\lambda) = 0$  with  $g'(\lambda) > 0$ . The range  $(\lambda_1, \lambda_2) = (\bar{\lambda}, \infty)$  of values of  $\lambda$  to consider can be divided into two sub-intervals  $(\bar{\lambda}, \lambda(t_0))$  and  $[\lambda(t_0), \infty)$  where respectively the case  $T_1 > t_0$  applies if  $g(\lambda(t_0)) > 0$  and the case  $T_1 = t_0$  applies if  $g(\lambda(t_0)) \leq 0$ . For the case  $T_1 > t_0$ , when  $g(\lambda(t_0)) > 0$  there exists a unique solution  $\lambda \in (\bar{\lambda}, \lambda(t_0))$  to  $g(\lambda) = 0$  since we can show that  $g(\bar{\lambda}) < 0$ . For the case  $T_1 = t_0$ , when  $g(\lambda(t_0)) \leq 0$  there exists a unique solution  $\lambda \in [\lambda(t_0), \infty)$  to  $g(\lambda) = 0$  since we can show that  $g(\infty) > 0$ .<sup>34</sup>

### Appendix B: Closed-form equations for the budget constraint

This appendix suggests a set of realistic assumptions to express the budget constraint as a series of closed-form equations. As is customary in this literature, the power utility function  $u(c) = c^{1-\gamma}/(1-\gamma)$ ,  $\gamma \neq 1$  is used. For the mortality and income assumptions, we opt for functional forms that are flexible enough to fit any discrete mortality table or income profile. For mortality, it is assumed that a constant force of mortality  $\mu_j$  applies at each age  $j$ . For the pre-retirement income function, we assume that the income process is continuous and that it grows at a rate  $g_j$  at age  $j$ . Let  $J(t) = j$  if  $j \leq t < j+1$ , the survival probability function for  $J(t) > t_0$  is given by  $p_{t_0, t} = e^{-\sum_{i=t_0}^{J(t)-1} \mu_i - \mu_{J(t)}(t - J(t))}$  and the pre-retirement income process by  $y_t = y_{J(t)} e^{g_j(t - J(t))}$ .

<sup>32</sup> More precisely,  $W_R(\lambda)$  is the accumulated value of  $s(c_i^*(\lambda))$  with interest over the interval  $[T_1(\lambda), R)$ . Setting  $W_R(\lambda)$  equal to the present value of  $s(c_i^*(\lambda_i))$  over the interval  $[R, T_2^i)$ , we can solve for  $\lambda_i$  with  $\lambda(T_2^{i-}) > \lambda_i \geq \lambda(T_2^i)$ . If there is only one period where  $\lambda(t)$  is decreasing after retirement, there is a unique solution to this equation. With multiple decreasing periods there can be more than one solution; in that case, the approach in Lachance (2012) can be used and the optimal one is the one with the highest  $\lambda_i$ .

<sup>33</sup> Lachance (2012) gives a detailed algorithm to handle the case with any number of periods where  $\lambda(t)$  is decreasing. This solution would apply before retirement until it remains only one period where  $\lambda(t)$  is decreasing.

<sup>34</sup> If  $\bar{\lambda} > \lambda(t_0)$ , only the case  $T_1 = t_0$  can apply and  $g(\bar{\lambda}) < 0$  guarantees the existence of a unique solution  $\lambda \in (\bar{\lambda}, \infty)$  to  $g(\lambda) = 0$ .



The budget constraint  $W(\lambda, \lambda_i) = 0$  in equation (A.5) is equal to the present value of savings over the interval  $[T_1, T_2]$ . From equations (13) and (4), savings are given by

$$s(c_i^*(\lambda)) = \begin{cases} \frac{y_t - \pi y_t - \tau_k(y_t - E) - G_k - c_i^*(\lambda)}{1 - \alpha\tau_k}, & t < R, \\ \frac{y_R + SS - \tau_k(y_R + M_h(y_R + SS/2) + H_{t,h} - E_R) - G_k - c_i^*(\lambda)}{1 - \alpha\tau_k(1 + M_h)}, & t \geq R. \end{cases} \quad (\text{B.1})$$

Note that in equation (B.1),  $c_i^*(\lambda)$  can take different forms: it can be the interior solution given in equation (22) or it can be  $C_k$  or  $C_h$  given in equations (14) and (15). Thus, the budget constraint is the sum of the present value of savings over a series of sub-intervals  $[t, T]$  where  $c_i^*(\lambda)$  takes the same form. For each interval  $[t, T]$ , the present value of savings can be expressed in closed form if we can integrate the present values of the functions  $y_t$ ,  $c_i^*(\lambda)$ ,  $H_{t,h}$ ,  $B_{t,h}^S$ , and some constants, which can easily be done as follows:

$$\int_t^T e^{-r(t-T_1)} dt = e^{rT_1} a_{t,T}^r, \quad \text{where } a_{t,T}^x = (e^{-xt} - e^{-xT})/x, \quad (\text{B.2})$$

$$\int_t^T e^{-r(t-T_1)} y_t dt = e^{rT_1} \sum_{j=J(t)}^{J(T)} e^{-g_{ij}} y_j a_{j,t,T}^{g_j-r}, \quad (\text{B.3})$$

where  $a_{j,t,T}^x = (e^{x\min(T,j+1)} - e^{x\max(t,j)})/x$ ,

$$\int_t^T e^{-r(t-T_1(\lambda))} c_i^*(\lambda) dt = e^{rT_1} \left( \frac{\lambda e^{(r-\beta)t_0}}{1 - \alpha\tau_k(1 + M_h)} \right)^{-1/\gamma} \times \sum_{j=J(t)}^{J(T)} (p_{t_0,j} e^{\mu_j})^{1/\gamma} a_{j,t,T}^{-r+(r-\beta-\mu_j)/\gamma}, \quad (\text{B.4})$$

$$\int_t^T e^{-r(t-T_1)} H_{t,h} dt = \begin{cases} e^{rT_1+ia} a_{t,T}^{r+i} H_{a,h}, & h = 1, 2, \\ 85\%SS \cdot e^{rT_1} a_{t,T}^r, & h = 3, \end{cases} \quad (\text{B.5})$$

$$\int_t^T e^{-r(t-T_1)} B_{t,h} dt = \begin{cases} e^{rT_1+ia} a_{t,T}^{r+i} B_{a,h}^S, & h = 1, 2, \\ SS \cdot e^{rT_1} a_{t,T}^r + e^{rT_1+ia} a_{t,T}^{r+i} (0.5X_a^1 + 0.35X_a^2)/0.85, & h = 3. \end{cases} \quad (\text{B.6})$$

These results can be combined with equations (14), (15), and (B.1) to obtain the present value of savings over an interval  $[t, T]$ . By multiplying them by  $e^{r(R-T_1)}$ , they can also be used to compute the retirement savings, tax liability, and tax subsidy measures in Section 7.

### Appendix C: Value function

The welfare gain is computed by first determining the percentage increase in consumption over the interval  $[T_1, T_2]$  that would make the individual indifferent between the Roth and traditional cases. With the exception of the interval  $[T_1, T_2]$ , this approach is equivalent to that presented in other works such as

Cocco *et al.* (2005).<sup>35</sup> To obtain a dollar measure, the percentage increase is multiplied by the value of consumption  $C_{T_1, T_2}^{\text{Roth}}$ . Let  $V_{T_1, T_2} = \int_{T_1}^{T_2} f(t)u(c_t^*)dt$  denote the value function, the welfare measure can be expressed as

$$\text{Welfare gain/loss with traditional account} = \frac{V_{T_1, T_2}^{\text{Trad } 1/(1-\gamma)} - V_{T_1, T_2}^{\text{Roth } 1/(1-\gamma)}}{V_{T_1, T_2}^{\text{Roth } 1/(1-\gamma)}} C_{T_1, T_2}^{\text{Roth}}. \tag{C.1}$$

<sup>35</sup> The binding periods are excluded from the calculation because including them biases the welfare measure. To see this, consider that by construction, the marginal utility of consuming an additional dollar in the binding period  $[t_0, T_1]$  is higher than in the non-binding period  $[T_1, T_2]$ . As a result, when the binding period  $[t_0, T_1]$  is included in the welfare calculation, a smaller increase in consumption is required to match a given welfare gain. Including the binding period  $[T_2, \omega]$  creates the opposite problem as marginal utility is lower in that period than in  $[T_1, T_2]$ .