Optimal Life-cycle Capital Taxation under Self-Control Problems*

Nicola Pavoni† and Hakki Yazici‡

August 11, 2015

Abstract

We study optimal taxation of savings in an economy where agents face self-control problems, and we allow the severity of self-control to change over the life cycle. We focus on quasi-hyperbolic discounting with constant elasticity of inter-temporal substitution utility functions and linear Markov equilibria. We derive explicit formulas for optimal taxes that implement the efficient (commitment) allocation. We show, analytically, that if agents' ability to self-control increases concavely with age, then savings should be subsidised and the subsidy should decrease with age. We also study the quantitative effects of age-dependent self-control problems and find that the optimal subsidies in our environment are much larger than those implied by models with constant self-control. Finally, we compare our optimal subsidies with those implied by the 401(k) plan. Although the subsidy levels in the two cases are of comparable magnitudes, the 401(k) plan implies an increasing pattern of subsidies while the optimal subsidies decrease over the life cycle.

JEL classification: E21, E62, D03.

Keywords: Self-control problems, Linear Markov equilibrium, Life-cycle taxation of savings.

*We thank the editor, Morten O. Ravn, two anonymous referees, as well as Per Krusell, John Leahy, and seminar participants at Bogazici University, Goethe University in Frankfurt, IIES, the SED meetings in Ghent, University of Alicante, University of Bologna, University of Oxford for their comments and suggestions. Yazici gratefully acknowledges financial support from the European Community Framework Programme through the Marie Curie International Reintegration Grant #268457.

†Bocconi University, IGIER, IFS, and CEPR. E-mail: nicola.pavoni@unibocconi.it

‡Sabanci University. E-mail: hakkiyazici@sabanciuniv.edu

Economists traditionally assume that people discount streams of utility over time exponentially. An important implication of exponential discounting is that under this assumption people have time-consistent inter-temporal preferences: How an individual feels about a given inter-temporal tradeoff is independent of when he is asked. However, laboratory and field studies on inter-temporal choice have cast doubt on this assumption.[1] This evidence suggests that discounting between two future dates gets steeper as we get closer to these dates. Such time-inconsistent inter-temporal preferences capture self-control problems. Naturally, all this evidence on self-control problems have led many economists to model this phenomenon and study its positive and normative implications.[2]

In this paper, we study optimal capital taxation over the life cycle in the presence of self-control problems. A common modelling assumption in the literature on self-control problems is that the degree of self-control problem is constant over time. This contrasts with the significant body of empirical research indicating that, like many other personality traits, people's ability to self-control changes as they age. Using an experimental approach to measuring self-control problems, one of the conclusions Ameriks *et al.* (2007) reach is summarised in the following quotation:

"A particularly interesting finding in Table 4 is the profound reduction in the scale of self-control problems as individuals age, which shows up only when one uses the absolute value of the self-control measure. Older individuals experience fewer self-control problems, either of overconsumption or under-consumption, than do their younger counterparts. This finding is certainly consistent with the common view that temptation falls with age, and has important connections with actual consumption behaviour over the life cycle. Models that allow for such a time-changing self-control parameter retirement may be necessary to explain the absence of a spike in consumption spending at the point when retirement assets become fully liquid."

Another set of evidence for changing levels of self-control problems comes from literature that investigates inter-temporal discounting over the life span. This research has shown that

[1]See DellaVigna (2009) for a survey of field studies and Frederick *et al.* (2002) for a survey of experimental studies. Also, see Laibson *et al.* (2007) for evidence of self-control problems in consumption asset holdings panel data.

[2]Three main models that have been proposed to capture self-control problems are the hyperbolic discounting model of Laibson (1997), the temptation model of Gul and Pesendorfer (2001), and the planner doer model of Thaler and Shefrin (1981).

the discrepancy between short term and long term discount rates falls with age predicting a life-cycle developmental trend toward increased self-control. Read and Read (2004) is an experimental study which estimates hyperbolic discount functions for three different groups of individuals: young adults, middle-aged, and old adults, with average ages of 25, 44, and 75. The authors find that the discrepancy between the short-term and the long-term discount rates is significantly higher for the young adults relative to the middle-aged. Similarly, this discrepancy is significantly higher for middle-aged adults compared to the old adults group, for which, they conclude the discrepancy virtually vanishes. Green *et al.* (1999) estimates hyperbolic discount functions for two groups of adults, young and old, with mean ages of 20 and 70 respectively, and finds that the discrepancy between short-term and long-term discount rates that exists for young adults disappears as people get old.[3]

A third set of evidence for changing self-control problems comes from personality psychology. As Ameriks *et al.* (2007) states "personality psychologists associate self-control with conscientiousness, one of the 'big five' personality factors."[4] There is a long list of empirical studies in personality psychology that show that conscientiousness, and in particular its lower-level facet, self-control, changes with age.[5] For example, in their survey article on personality development in adulthood, Caspi *et al.* (2003) conclude that: "it appears that the increase in conscientiousness is one of the most robust patterns in personality development, especially in young adulthood."

In this paper, we extend the traditional models of self-control to allow for varying degrees of self-control problem over the life cycle, and study optimal capital taxation. In our model, agents make consumption and savings decisions facing self-control problems at all ages. In the last period of their lives, people make consumption and bequest decisions knowing that they are going to be replaced by their offspring next period. We model preferences that exhibit self-control problems through the quasi-hyperbolic discounting framework of Laibson (1997),

[3]For more empirical work on life-span analysis of inter-temporal discounting, see also Green *et al.* (1994) and Green *et al.* (1996).

[4]Actually, Ameriks *et al.* (2007) validates this relationship between conscientiousness and the measure of self-control used in the experiment (the EI gap) and finds that "the data reveal a strong relationship between the conscientiousness questions and the absolute value of the EI gap."

[5]For example, see John *et al.* (2003) and Helson *et al.* (2002). Ameriks *et al.* (2007) also, through their experimental finding, show that there is a profound reduction in the scale of self-control and conscientiousness problems as individuals age.

which builds on the seminal works of Strotz (1955) and Phelps and Pollak (1968). We extend the Laibson (1997) model by allowing for the degree of self-control problem to change over time. We assume people are sophisticated, meaning they anticipate future self-control problems.

In this environment, we define *efficient (or commitment)* allocation as the allocation that would arise in the absence of self-control problems. It is given by the solution to a fictitious social planner's consumption-saving problem where the planner future utilities *exponentially*. In our environment, this preference corresponds to the preference of an initial generation parent. The main exercise in this paper is to examine the optimal tax policy that implements the efficient allocation. In this sense, this paper is a normative exploration of optimal paternalistic tax policy regarding life-cycle saving behaviour. It is well-known that in models of quasi-hyperbolic discounting there is multiplicity of equilibria.[6] We restrict attention to the (unique) linear Markov equilibrium of our economy.

We derive closed form formulas for optimal age-dependent capital taxes. Our closed form solution represents the equilibrium obtainable as the limit of the equilibria of finite-period economies. We show that optimal capital taxes can be positive as well as negative in different periods of life and they can be increasing, decreasing, or changing non-monotonically with age, depending on what we assume about the evolution of self-control problem over the life cycle. This ambiguity result about the qualitative properties of optimal taxes shows that researchers who take self-control problems seriously should also take the evolution of self-control problems over the life cycle seriously before making policy suggestions. This result also questions the basic presumption in the literature that self-control problems always imply optimality of saving subsidies, which, as we show, arises from the assumption of constant self-control over age.

Our closed form tax formulas are obtained assuming agents have constant elasticity of intertemporal substitution (CEIS) preferences. The formulas are valid (i) if CEIS coefficient is one (the utility function is logarithmic) or (ii) for any CEIS coefficient if the economy is in a steady-state. Using these formulas, we prove that if, as suggested by the available empirical evidence, the degree of self-control increases with age, then capital should indeed be *subsidised* in all periods. We also prove that, if self-control increases concavely with age, then optimal capital subsidies should *decrease* with age.

We study the quantitative effects of age-dependent self-control problems in a calibrated ver-

[6]For discussions of multiplicity of equilibria, see, among others, Laibson (1994) and Krusell and Smith (2003).

sion of our model and find that the optimal subsidies in our model with decreasing self-control problems are *much larger than those implied by a model with constant self-control problems*. We also compare our optimal subsidies with those implied by the 401(k) plan. If we exclude the very last periods before retirement - where the subsidy rate in the 401(k) essentially mimics the employer matching rate - the subsidy levels in the two cases are of comparable magnitudes. A marked difference emerges however: the 401(k) plan implies an increasing pattern of subsidies while the optimal subsidies decrease over the life cycle.

Our benchmark model assumes perfectly altruistic parents, making it equivalent to a standard infinite horizon framework. Section 3.1 allows for imperfect altruism and generalises our optimal tax formulas to take into account the finite life time effects of Krusell *et al.* (2010). We find that, in our model, the effect of age-dependent self-control dominates the finite time effect induced by imperfect altruism: optimal capital subsidies decrease over the life cycle even when parents do not care at all about their offsprings (i.e., when finite life time effect is the strongest). In Section 3.2, we extend the finite horizon version of the model developed in Section 3.1 with borrowing constraints to study the life-cycle consumption implications of our model. We find that, even though our model abstracts from important life cycle issues such as child-rearing and health, life-cycle consumption profiles implied by our calibrated self-control patterns capture key features of empirical life-cycle consumption profiles fairly well.

Related Literature. Our paper is related to a number of recent papers that have explored the implications of self-control problems for optimal paternalistic taxation. O'Donoghue and Rabin (2003) and O'Donoghue and Rabin (2006) analyse models of paternalistic taxation of unhealthy goods. More closely related is Krusell *et al.* (2010), which analyses optimal taxation of savings in an economy where agents live finitely many periods and have self-control problems à la Gul and Pesendorfer (2001).[7] First, they prove that the optimal policy prescriptions of the quasi-hyperbolic model and the temptation model are identical when the utility function is logarithmic or when it is CEIS and the temptation parameter goes to infinity. Second, they show that savings should be subsidised and that this *subsidy should be increasing with time* due to finite life time effect.[8] Our work differs from this paper most importantly by allowing for chang-

[7]Krusell *et al.* (2002) also analyse optimal taxation of savings under self-control problems but their main focus is on an environment where the government as well as the people face issues of time-inconsistency.

[8]It is indeed straightforward to show that, in the infinite horizon version of their model, the subsidies would be constant.

ing levels of self-control problems over the life cycle: like all papers prior to ours, Krusell *et al.* (2010) assume that the degree of self-control problem is constant over time. The implications of modelling age-dependent self-control problems turn out to be significant. First, by assuming empirically plausible patterns of self-control problems over the life cycle, we show, analytically, that capital subsidies should actually be decreasing with age. Section 3.1 shows that this result continues to hold even when we extend the model to take into account the finite life time effects of Krusell *et al.* (2010). Second, we find that, for an agent with a given level of self-control problem, the age-dependence of self-control model imply much higher levels of optimal subsidies relative to the ones implied by the constant self-control model.

Another important paper that is related to ours is Imrohoroglu *et al.* (2003) which studies the role of social security in a model where agents have self-control problems. They consider a rich overlapping generations model with uninsurable unemployment shocks and liquidity constraints, and find that social security is not very useful in helping agents solve their self-control problems. Ours is a theory of capital subsidies under complete markets. One advantage of our analysis is that, whenever utility is logarithmic, our results are robust to many dimensions of heterogeneity - such as the life-cycle wage profile and the wealth distribution - whereas the normative predictions in models with incomplete markets may obviously depend on all these features.

As discussed above, an immediate implication of age-dependent self-control problems is that capital taxes should be age-dependent. The age-dependence result is also a feature of two sets of earlier contributions that analyse benefits of age-dependent capital income taxes with time-consistent agents. First, in the Ramsey taxation tradition, Erosa and Gervais (2002) shows that, in life-cycle economies, if the government has access to age-dependent linear capital and labor income taxes, the resulting optimal tax system features age-dependence both for capital and labor income. Second, the New Dynamic Public Finance literature calls for age-dependence in optimal capital and labor income tax codes (e.g., Farhi and Werning (2013) and Golosov *et al.* (2011)). The forces generating age-dependence in the current paper, however, are completely different from the forces in these papers.[9] Therefore, our paper complements this literature by

[9]The optimality of age-dependence in Erosa and Gervais (2002) is a direct implication of time-dependent consumption and labor plans that is valid even in the steady state due to life cycle changes in people's productivity. In the New Dynamic Public Finance models, capital is taxed in order to deter people from joint deviation of saving

providing a new mechanism through which capital taxes should depend on age. As such, in this paper, the life-cycle pattern of optimal capital taxes depends on features of the environment that are neglected by these papers.

The rest of the paper is organised as follows. Section 1 lays down the basiline model whereas Section 2 analyses optimal subsidies both qualitatively and quantitatively. In Section 3, we consider imperfect altruism extension and the life-cycle consumption implications of of the calibrated self-control model. Section 4 concludes. The proofs of propositions, the procedure we employ to approximate quasi-hyperbolic discount functions, and the analysis of optimal subsidies under partial sophistication and in the presence of an illiquid asset are presented in a separate online Appendix.

1 Model

The economy is populated by a continuum of a unit measure of dynasties who live for a countable infinity of periods, $t = 1, 2, \ldots$, where each agent within a dynasty is active for $I + 1$ periods. In the first I periods, agents make consumption saving decisions facing different degrees of self-control problems at different ages. In the last period of their lives, agents decide how much to consume and bequeath to their offspring, who replace them in next period. People are altruistic and they anticipate their offspring's self-control problems.We use quasi-hyperbolic discounting framework formalised by Laibson (1997) to model self-control problems as follows. An agent in his ultimate period of life (we refer to this agent as the parent hereafter) has the following preferences over dynastic consumption stream:

$$u(c_0) + \delta u(c_1) + \delta^2 u(c_2) + \cdots + \delta^I u(c_I) + \delta^{I+1} u(c_0') + \ldots$$

where c_0, c_i, and c_0' refer to the the consumption levels of the current parent, the offspring at age i, and the offspring when he becomes a parent, respectively. u is the instantaneous utility function and δ refers to both the discount factor and the altruism factor. The offspring has

and shirking. Since people at different ages (and contingencies) have different levels of accumulated wealth and future prospects, they have different tendencies to save, and hence, the corrective taxes depend on age.

different preferences at different ages. His preferences are given by

$$u(c_i) + \beta_i \delta \left[u(c_{i+1}) + \delta u(c_{i+2}) + \cdots + \delta^{I-i} u(c'_0) + \ldots \right], \text{ at age } i \in \{1, ..., I-1\},$$
$$u(c_I) + \beta_I \delta \left[u(c'_0) + \ldots \right], \text{ at age } I.$$

When $\beta_i = 1$ for all i, all agents at all ages are time-consistent as there is no self-control problem. Throughout the paper we will assume that $\beta_i < 1$, meaning individuals postpone their planned savings when the date of saving comes. Prior to the current paper, the literature has assumed $\beta_i = \beta$ for all i, meaning the degree of self-control problem is constant as people age. Following the large body of empirical findings provided by personality psychologists and experimental studies, we allow for the severity of self-control problems, β_i to depend on i. In the baseline model, we assume people are fully sophisticated, meaning they fully anticipate the self-control problems faced by future selves and descendants. In Appendix C, we analyse a version of the model in which people are only partially aware of their self-control problems. [10]

The instantaneous utility function, u, is of the CEIS form with elasticity parameter $\sigma > 0$:

$$\begin{aligned} u(c) &= \frac{c^{1-\sigma}}{1-\sigma}, \text{ for } \sigma \neq 1; \\ &= \log(c), \quad \text{else.} \end{aligned}$$

Production takes place at the aggregate level according to the function $F(k, l)$, where k is aggregate capital and l is aggregate labor. The production function satisfies the usual neoclassical properties together with the Inada conditions:

$$F_1, F_2 > 0 \ ; \ F_{11}, F_{22} \leq 0; \quad \text{and} \quad \lim_{k \to 0} F_1 = \infty; \quad \lim_{k \to \infty} F_1 = 0.$$

Labor is inelastically supplied, so at all dates $l = 1$. Define $f(k) = F(k, 1) + (1-d)k$, where d refers to the fraction of capital that is forgone due to depreciation. There is a credit market in which agents can trade one period risk-free bonds and capital as perfectly substitutable assets. Since at any given date there is not cross-sectional heterogeneity, all agents have the same level of asset holdings. Hence, letting b_t be the amount of asset holdings of the agent alive in period t, the credit market clearing condition is $k_t = b_t$.

[10] See Ariely and Wertenbroch (2002) for behavioural evidence on partial sophistication.

1.1 The Efficient Allocation

The *efficient* or – as we use interchangeably throughout the paper – the *commitment* allocation is the allocation that would arise in the absence of self-control problems. It is given by the solution to a fictitious social planner's consumption-saving problem where the planner discounts exponentially with discount factor δ. In our environment, this preference corresponds to the preference of an initial generation parent. By taking a long-term perspective and evaluating welfare according to the initial parent's preference, we are following much of the literature.[11] The following Euler Equations characterise the efficient allocation, which we denote with an asterisk throughout the paper:[12]

$$\begin{aligned} u'(c_i^*) &= \delta f'(k_i^*)u'(c_{i+1}^*), \text{ for } i \in \{0, 1, \ldots, I-1\}, \text{ and} \\ u'(c_I^*) &= \delta f'(k_I^*)u'(c_0'^*), \ldots \end{aligned} \tag{1}$$

1.2 Implementing the Efficient Allocation

Since people in this economy face self-control problems, laissez-faire market equilibrium cannot attain the commitment allocation. Our main interest in this paper is to find and characterise a tax system that implements the commitment allocation in the market environment. We call such a tax system *optimal.* We proceed by defining a market equilibrium with taxes. It is important to note that from the outset we restrict the set of taxes that are available to the government to linear taxes on savings coupled with lump-sum rebates (throughout the paper we call this the set of linear taxes). In general, it is not obvious that there is a linear tax system that implements the efficient allocation. However, since we focus our attention to linear equilibria, a linear tax system that implements the efficient allocation exists. We will verify this claim in Section 2.

1.3 Markov Equilibrium with Taxes

For notational simplicity, here in the main text, we only present the stationary version of the model where the level of aggregate capital stock starts from its steady-state level, k. The prices

[11] See DellaVigna and Malmendier (2004), Gruber and Koszegi (2004) and O'Donoghue and Rabin (2006), for example.

[12] We do not state the transversality condition but the commitment allocation will converge to a steady state with positive capital as long as $k_0 > 0$.

at the steady state are given by

$$R = f'(k) \quad \text{and} \quad w = f(k) - f'(k)k. \tag{2}$$

In such a world, the only index we need to carry around is the age index i. In Appendix A.1, we provide the general setup where the economy starts from an arbitrary level of capital stock and prices change over time. We prove our main result, Proposition 1, for the general case, and show that if the utility function is logarithmic, then optimal taxes do not depend on whether the economy is at a steady state or in a transition.

Let τ_i be the savings (capital) tax agent $i = 0, 1, \ldots, I$ pays. Tax proceeds are rebated in a lump-sum manner in every period. Denote the lump-sum rebate in period i by T_i and let $\tau = \{\tau_i, T_i\}_{i=0}^{I}$. For each set of taxes, we define the policy functions $b_i(\cdot; \tau)$ for $i = 0, 1, \ldots, I$, describing the optimal level of asset holdings of agent i given prices and taxes. When agent of age n is deciding b_n, his evaluation of the effect of his choice on b_i, $i > n$ will be described by the (nested) function $b_i(b_{i-1}(...b_{n+1}(b_n; \tau)...; \tau); \tau)$, which will be referred to as $b_i(...(b_n)...))$ so as to simplify notation. In addition, in order to only deal with functions, we assume each agent's solution is unique, a property satisfied by our closed form solution involving linear policies. Of course, in case of multiple solutions, our policy functions correspond to appropriate selections from the policy correspondences.

A Stationary Markov equilibrium *with taxes τ consists of a level of capital k, prices R, w, value function $V(\cdot; \tau)$ and policy functions $\{b_i(\cdot; \tau)\}_{i=0}^{I}$ such that: (i) the prices satisfy (2); (ii) the value functions and the policies are consistent with the parent's problem described below; (iii) the government budget is satisfied period-by-period and markets clear: $T_i = R\tau_i b_i(k; \tau)$ and $b_i(k; \tau) = k$ for all i.*

We now formally define the parent's problem. Let $V(b; \tau)$ be the value of a parent's problem who had b units of assets in his last period before parenthood and faces the tax system τ. The parent chooses his bequest b_0 and does not have any direct control over $b_1, \ldots, b_I$. Note that his preferences are not aligned with his offspring's (in a given period i, parent's discount factor is δ whereas offspring's is $\beta_i \delta$). The parent foresees this misalignment of preferences, and correctly forecasts future policies.

The parent's value and the policies are given by the solution to the following problem:

$$
\begin{aligned}
V(b;\tau) = \max_{b_0} \; & u\left(R\left(1-\tau_I\right)b + w + T_I - b_0\right) \\
& +\delta\left\{\sum_{i=0}^{I-1}\delta^i u\left(R\left(1-\tau_i\right)b_i(...(b_0)...) + w + T_i - b_{i+1}(...(b_0)...)\right) + \delta^I V\left(b_I(...(b_0)...);\tau\right)\right\} \\
\text{s.t. for all } b_0 & \\
b_1(b_0;\tau) = \arg\max_{\hat{b}_1} \; & u\left(R\left(1-\tau_0\right)b_0 + w + T_0 - \hat{b}_1\right) + \delta\beta_1 u\left(R\left(1-\tau_1\right)\hat{b}_1 + w + T_1 - b_2(\hat{b}_1)\right) \\
& +\delta\beta_1\left\{\sum_{i=2}^{I-1}\delta^{i-1}u\left(R\left(1-\tau_i\right)b_i(...(\hat{b}_1)...) + w + T_i - b_{i+1}(...(\hat{b}_1)...)\right) + \delta^{I-1}V\left(b_I(...(\hat{b}_1)...);\tau\right)\right\} \\
\text{s.t. for all } b_1 & \\
b_2(b_1;\tau) = \arg\max_{\hat{b}_2} \; & u\left(R\left(1-\tau_1\right)b_1 + w + T_1 - \hat{b}_2\right) + \delta\beta_2 u\left(R\left(1-\tau_2\right)\hat{b}_2 + w + T_2 - b_3(\hat{b}_2)\right) \\
& +\delta\beta_2\left\{\sum_{i=3}^{I-1}\delta^{i-2}u\left(R\left(1-\tau_i\right)b_i(...(\hat{b}_2)...) + w + T_i - b_{i+1}(...(\hat{b}_2)...)\right) + \delta^{I-2}V\left(b_I(...(\hat{b}_2)...);\tau\right)\right\} \\
\ldots & \\
b_{I-1}(b_{I-2};\tau) = \arg\max_{\hat{b}_{I-1}} \; & u\left(R\left(1-\tau_{I-2}\right)b_{I-2} + w + T_{I-2} - \hat{b}_{I-1}\right) \\
& +\delta\beta_{I-1}\left\{u\left(R\left(1-\tau_{I-1}\right)\hat{b}_{I-1} + w + T_{I-1} - b_I(\hat{b}_{I-1})\right) + \delta V\left(b_I(\hat{b}_{I-1});\tau\right)\right\} \qquad (3) \\
\text{s.t. for all } b_{I-1} & \\
b_I(b_{I-1};\tau) = \arg\max_{\hat{b}_I} \; & u\left(R\left(1-\tau_{I-1}\right)b_{I-1} + w + T_{I-1} - \hat{b}_I\right) + \delta\beta_I V\left(\hat{b}_I;\tau\right). \qquad (4)
\end{aligned}
$$

To understand the nested nature of policies better, let us analyse the definition of policies in (3) and (4). First, constraint (4) describes how self I chooses b_I. This agent chooses b_I anticipating correctly that next period when he becomes a parent he will face an offspring with self-control problems, and the offspring will face an offspring with self-control problems, and so on. Second, consider constraint (3) which defines how self $I-1$ chooses b_{I-1}. Being sophisticated, self $I-1$ knows that his followers will have self-control problems. In particular, self $I-1$ knows that self I chooses b_I according to (4). We have just seen that the last constraint, (4), enters the parent's problem in at least two ways: first, in the definition of self I's policy function and then as a constraint in the definition of self $I-1$'s policy function. These two different constraints are represented by a single constraint, (4), because the parent and self $I-1's$ sophisticated belief agree about how self I will behave. Similarly, the constraint describing self $I-1$'s policy is also a constraint in the constraint that describes self $I-2$'s policy, and self

$I-2$'s policy is also a constraint of self $I-3$'s, and so on. Thus, actually the constraint that describes the policy of self i enters parent's problem in i different places but since these are all identical constraints, we represent them with just one constraint that describes self $i's$ policy.

We restrict attention to linear equilibria, meaning equilibria in which policy functions are linear in net present value of current wealth. Mathematically, defining $\Gamma_{i-1}(b;\tau)$ to be the net present value of wealth available to an agent at the beginning of age i with asset level b and under tax system τ, we derive closed form solutions of the form $c_i(b;\tau) = M_i(\tau)\Gamma_{i-1}(b;\tau)$.

2 Optimal Taxes

In this section we analyse optimal capital taxes in the model introduced in Section 1. Proposition 1 characterises optimal taxes when utility is logarithmic.

Proposition 1 *Suppose $u(c) = \log(c)$. The optimal taxes are given by:*

$$1 - \tau_i^* = \frac{1}{\beta_i}\left(1 - \delta + \beta_{i+1}\delta\right), \textit{ for } i \in \{0, 1, \ldots, I\},$$

with $\beta_0 = \beta_{I+1} = 1$.

Proof. Relegated to Appendix A.1. ■

The main task in proving Proposition 1 is to compute linear policies, $M_i(\tau)$, which we do as follows. The assumption of linearity of policy functions implies that offspring's problems at all ages are strictly concave maximisation problems. This implies that the first-order optimality conditions are not only necessary but also sufficient. As a result, in the parent's problem, we can replace the constraints that take the form of maximisation problems with the corresponding first-order optimality conditions. Using the the first-order approach and a version of a guess and verify method, we find analytic expressions for the value function V and the vector of constants $M_i(\tau)$ describing equilibrium linear policies. Then, we use the policy functions to compute Euler equations that describe people's optimality conditions regarding savings at different ages in equilibrium. The comparison of these equilibrium Euler equations with the planner's Euler equations, given by (1), gives the optimal distortions (taxes) that need to be created to implement the commitment allocation.

It might be important to stress that - as shown in Appendix A.1. - Proposition 1 holds regardless of whether the economy is in a steady state or in a transition. In particular, since agents do not face binding liquidity constraints, *the expressions for taxes hold for any life-cycle path of wages*. In Appendix C, we also prove that, under the logarithmic utility assumption, the optimal taxes given by Proposition 1 are valid independent of whether we assume people are aware of their future self-control problems or not.

Proposition 2 shows that if the economy is in a steady state, then the optimal taxes characterised in Proposition 1 for the $\sigma = 1$ case are valid for any $\sigma > 0$.

Proposition 2 *Assume k is such that $\delta f'(k) = 1$. Then, for any $\sigma > 0$, optimal taxes are given by the exact same expressions as in Proposition 1.*

Proof. Relegated to Appendix A.2. ■

The tax formula for in Proposition 1 consists of two components. The first component, $\frac{1}{\beta_i}$, is easier to understand. Because of his *current* self-control problem, self i discounts tomorrow by an extra β_i and hence wants to under-save relative to the efficient allocation. By factoring the after tax return with $\frac{1}{\beta_i}$, we can exactly offset the extra discounting, thereby getting rid of this under-saving motive. We call this first part of the tax formula the *current component*. Clearly, the current component is always greater than one, i.e., it always calls for a subsidy.

The second component of the tax formula, $(1 - \delta + \beta_{i+1}\delta)$, is there to correct deviations in current savings caused by suboptimal actions of future selves. This part the tax formula, which we refer to as the *future component*, is always less than one, meaning it calls for a tax on savings. Intuitively, from self i's perspective, self i+1 is under-saving due to his self-control problems ($\beta_{i+1} < 1$.) Thus, self i's welfare would go up if he can make self i+1 increase his savings, which self i can achieve by increasing his own savings since, under our parametric assumptions, self i+1's savings is strictly increasing in self i's savings. This discussion implies that self i has an additional marginal benefit of saving in equilibrium when $\beta_{i+1} < 1$ relative to the case where $\beta_{i+1} = 1$. It is this extra benefit that makes self i save more relative to the commitment level. The future component of the tax formula calls for a tax in order to offset this extra return, and ensure self i does not over-save. The reader might still feel puzzled by our argument: after all, when facing the optimal taxes, self $i + 1$ saves the efficient amount. Notice, however, that from self i's perspective, self $i + 1$ is still under-saving (at the new price that is inflated by the subsidy), and hence future component of the tax is still needed.

Obviously, the sign of the optimal capital tax depends on whether the current or the future component dominates. For constant self-control (i.e., $\beta_i \equiv \beta < 1$), the current component always dominates, implying an optimal negative tax (i.e., optimality calls for a saving subsidy). We will see below that when β_i changes with age, depending on the pattern of change, either component may dominate, and the optimal tax can in general be positive or negative.

Notice that τ_0^* is only shaped by the future component. It is hence always positive. Since it is applied to the wealth transferred to future generations, τ_0^* can be interpreted as a bequest tax. In this paper, we do not analyse taxation of wealth transferred across generations. We study this topic in detail in Pavoni and Yazici (2013b).

2.1 Lessons for Capital Taxation

Propositions 1 and 2 imply several general lessons for capital taxes which are summarised below in a series of corollaries.

Corollary 3 ***(Age-dependence)*** *Optimal capital taxes are age-dependent. In particular, depending on how the degree of self-control changes with age: (i) Optimal capital taxes might be positive or negative at different ages. (ii) Optimal capital taxes might be increasing or decreasing with age at different ages.*

Proof. (i) For an example of $\tau_i > 0$, set $\beta_{i+1} \approx 0$ and $\beta_i > 1 - \delta$. For an example of $\tau_i < 0$, set $\beta_i = \beta_{i+1} = \beta < 1$. (ii) See the brown line with crosses in Figure 1 for an example. ■

The optimal capital taxes depend on age as long as the degree of self-control problems depend on age. Figure 1 plots various life cycle profiles of self-control on the left panel and the corresponding optimal taxes on the right panel. The figure shows that, under constant self-control of $\beta = 0.5$, depicted by blue dots, the subsidy is also constant over the life cycle at 4%. Corollary 3 also states that, unlike the common presumption in the literature, it might be optimal to tax people with self-control problems. However, although a theoretical possibility, optimal capital taxes can be positive in our model only under parameter specifications that seem to be inconsistent with data. Figure 1 shows that the model can generate a few periods of positive optimal capital taxes only when β_i declines sharply with age, as depicted by the brown crosses. As we discuss in Section 2.2, increasing level of self-control problems with age (β_i declining with i) is at odds with empirical findings. The orange dashed line and the light blue solid line in Figure 1 display self-control patterns that are increasing with age. In both cases, capital should be subsidised. Then, whether subsidies should increase or decrease with age depends on the curvature of β_i.

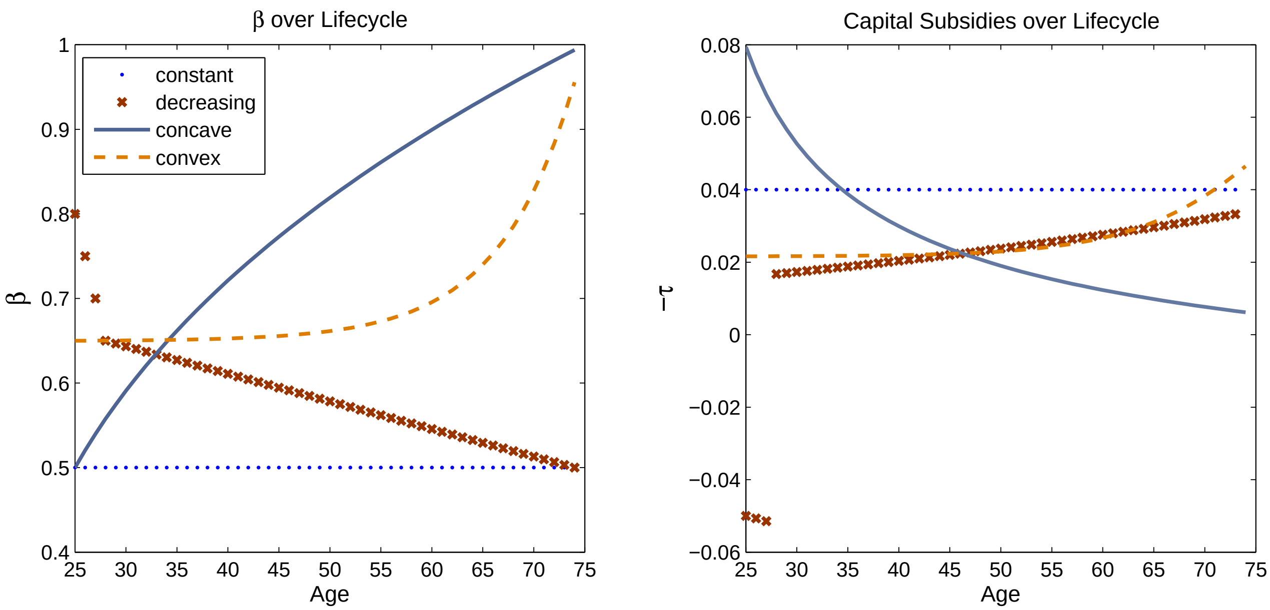


Figure 1: *Left panel: Examples of constant, decreasing, concave increasing and convex increasing patterns self-control problems over the life cycle. Right panel: Corresponding optimal capital subsidies.*

Corollaries 4 and 5 below characterise quite sharply the sign and the monotonicity properties of optimal capital taxes over the life cycle under the assumption that self-control problems decrease concavely with age (β_i is increasing and concave in i). This pattern is in line with the available empirical evidence and agrees with the self-control profiles we calibrate in Section 2.2.[13] Corollary 4 shows that if the severity of self-control problems decline with age, then

[13]Several different strands of literature provide evidence on this pattern of self-control problems. First, research on inter-temporal discounting over the life span has shown that short term discount rates fall with age predicting a life-cycle developmental trend toward increased self-control. See, in particular, Green *et al.* (1999), and Read and Read (2004). Second,Ameriks *et al.* (2007) finds that 'EI gap,' the measure of self-control problem used in the paper's experiment, decreases concavely with age. Finally, personality psychologists associate self-control with conscientiousness, one of the 'big five' personality factors, and in the words of Caspi *et al.* (2003) 'it appears that the increase in conscientiousness is one of the most robust patterns in personality development, especially in young adulthood.' John *et al.* (2003) estimates conscientiousness as a quadratic function of age and finds that the quadratic age term has a negative coefficient 'indicating that the rate of increase [in conscientiousness] was greater at younger ages than at older ages.' Roberts *et al.* (2006) also estimates a concave conscientiousness

capital should be subsidised at all ages.

Corollary 4 ***(Optimality of Capital Subsidies)*** *If $\beta_{i+1} \geq \beta_i$, for all $i \geq 1$, optimal capital tax is negative for all ages.*

Proof. $1 - \tau_i^* = \frac{1}{\beta_i}(1 - \delta + \beta_{i+1}\delta) > \frac{\beta_{i+1}}{\beta_i} \geq 1$. ■

Corollary 5 shows that, if people's ability to self-control increases concavely with age, then capital subsidies should decrease with age.

Corollary 5 ***(Decreasing Capital Subsidies)*** *If $0 \leq \beta_{i+1} - \beta_i \leq \beta_i - \beta_{i-1}$ for all $i \geq 1$ (concavity), then optimal capital subsidies decrease with age.*

Proof. $1 - \tau_{i-1}^* = \frac{1-\delta}{\beta_{i-1}} + \frac{\beta_i \delta}{\beta_{i-1}} > \frac{1-\delta}{\beta_i} + \frac{\beta_i \delta}{\beta_{i-1}} > \frac{1-\delta}{\beta_i} + \frac{\beta_{i+1}\delta}{\beta_i} = 1 - \tau_i^*$, where the first and second inequalities follow from $\beta_{i-1} < \beta_i$ and $\beta_{i+1} - \beta_i \leq \beta_i - \beta_{i-1}$, respectively. ■

The result of Corollary 5 differs from Krusell *et al.* (2010) which concludes that in any finite economy with constant self-control, capital subsidies should be increasing with age. The optimality of increasing subsidies in their case is due to the finite life time people face, and this element is missing from our analysis due to our assumption of perfect altruism. In Section 3.1, we show that the finite life time effect is quantitatively small within the relevant parameter space, implying that the optimality of decreasing subsidies with age is generally optimal.

2.2 Quantitative Analysis

In this section, we quantitatively analyse optimal capital taxation over the life cycle assuming either one of the justifications of the tax formulas in Proposition 1 hold: either utility is logarithmic or the economy is at a steady state. In order to conduct a numerical analysis, we have to choose values for the parameters of the model. Individuals are assumed to be born at the real time age of 25 and they live 51 years, so they die at the end of age 75. Observe that the tax formulas do not depend on the constant relative risk aversion coefficient σ, the shape of the production function F, or the depreciation rate, d. Therefore, we do not specify values for

pattern.

these parameters.[14] The only parameter values needed for the analysis are the true yearly discount factor δ and the evolution of self-control with age, $\{\beta_i\}_{i=1}^{51}$. Here, β_i represents the level of self-control problem at real age of $i+24$.

We set the long-run yearly discount factor $\delta = 0.96$. This value corresponds to the benchmark estimate in a constant self-control model with $\sigma = 1$ by Laibson *et al.* (2007). The self-control vector, $\{\beta_i\}_{i=1}^{51}$, is calibrated as follows. We assume the relationship $i \to \beta_i$ takes the following functional form:

$$\beta_i = a - d \exp\left\{\frac{51-i}{b}\right\}. \tag{5}$$

An advantage of this functional form over some other - perhaps simpler - forms is that it is relatively easy to ensure that it satisfies a key condition of our model, namely $\beta_i \leq 1$ for all $i \in \{1, ..., 51\}$.[15] This form is also flexible in the sense that it allows for both concave and convex and decreasing and increasing patterns of self-control over the life cycle. This flexibility is important as we do not want to - a priori - put any restrictions on the self-control pattern. We calibrate the parameters of the self-control function in (5) in two alternative ways.

RR Calibration. Read and Read (2004) conducts a survey with 129 respondents between the ages of 19 and 89 in which respondents are asked to make a large number of time discounting decisions on both computerised and paper-and-pencil questions. The study estimates hyperbolic discount functions for three age groups: young, middle-aged, and older adults, with mean ages of 25, 44, and 75, respectively. Unlike quasi-hyperbolic discount functions, which have one short-term and one long-term discount factors (β and δ), hyperbolic discount functions allow many different discount factors depending on the length of delay for future reward. This implies that the hyperbolic discount functions estimated in Read and Read (2004) are not readily available for our study. Therefore, our calibration strategy works in two stages. First, for each age group, we find the β that best approximates the hyperbolic discount function for that group. We assume that the β at the mean age of a group is equal to the β that is approximated for that group (for example, β at age 25 equals β of young adults). Second, we use the β for ages 25, 44,

[14]Observe that if we want the taxes computed using the formulas in Proposition 1 to be valid under any σ, then we need to assure that the interest rate R or the deeper parameters of the production function F and d satisfy the steady-state condition $R = f'(k) = F'(k,1) - d = \delta^{-1}$, where k refers to steady-state level of capital stock.

[15]To be precise, β_i is bounded above by 1 with this functional form under the assumption that $a = 1 + d$. As we see when we discuss our calibration exercises in the next section, this assumption is empirically supported.

and 75 (or correspondingly for periods $i = 1, 20$, and 51) to pin down the parameters a, b, and d of the functional form in (5).

For the older adults group, the approximation stage is quite simple because Read and Read (2004) finds that the older adults group display exponential discounting. The authors write: "Green et al's major result- that younger people show hyperbolic discounting while older people show exponential discounting - is supported by our data." This implies that a single (long-term) discount factor, δ, is sufficient to describe the behaviour of older adults, implying this group has $\beta = 1$. Since age 75 corresponds to period 51 in our period, this gives us our first calibration target, $\beta_{51} = 1$, which implies $a = 1 + d$ when plugged in (5).

For the young adults and middle-aged groups, the approximation procedure chooses β using a least squares procedure: that is, β is chosen to minimise the sum of the squares of errors between the yearly discount factors that are implied by the hyperbolic discount function and our (δ, β) model. The details of our approximation procedure can be found in Appendix B. The $\beta's$ that come out of our approximation procedure are 0.525 and 0.732 for the age 25 and age 44 groups, respectively. Thus, we set $\beta_1 = 0.525$ and $\beta_{20} = 0.732$ as the two other calibration targets. These two targets allow us to identify the remaining parameters of the self-control function (5), b and d. The calibration targets and the calibrated parameter values are reported in the upper panel of Table 1.

GMO Calibration. This calibration uses a strategy that is quite similar to RR calibration except that the calibration targets come from two other studies. Two of the targets of GMO calibration come from Green *et al.* (1999) which is an earlier paper that also uses an experimental approach to estimate hyperbolic discount functions. They do so for two groups of adults: the young and the old, with average ages of 20 and 70, respectively. Similar to Read and Read (2004), Green *et al.* (1999) also finds that the inter-temporal discounting behaviour of their old adults group can be best described by exponential discounting function. Therefore, we set β for this group to be 1. Using the exact same approximation procedure we use in RR calibration, we find β for the age 20 group to be 0.362. Then, we identify (5) for the 51 periods starting at age 20 and ending at age 70. Thus, we set $\beta_1 = 0.362$ and $\beta_{51} = 1$. The latter again implies that $a = 1 + d$ as in the RR calibration. We have two parameters, b and d to be identified and only one target $\beta_1 = 0.362$. We still need one more calibration target to identify b and d. We use the average

RR Calibration

Parameters	Targets	Source
d=0.795	β at age 25 $= 0.525$	RR
b=106.74	β at age 44 $= 0.732$	RR
a=1+d	β at age 75 $= 1$	RR

GMO Calibration

Parameters	Targets	Source
d=0.055	β at age 20 $= 0.362$	GMO
b=19.73	Average $\beta = 0.818$	LRT
a=1+d	β at age 70 $= 1$	GMO

Table 1: *This table reports the two alternative calibration exercises that we conduct for self-control pattern over the life cycle. The acronyms GMO, LRT, and RR stand for Green* et al. *(1999), Laibson* et al. *(2007), and Read and Read (2004), respectively. a, b and d are the parameters of the self-control function in expression (5).*

level of self-control problems in the economy, call it β_{avg}, as an additional target. We take β_{avg} to be 0.818 in our benchmark analysis following the estimate of Laibson *et al.* (2007) for $\sigma = 1$ for a constant self-control model. The calibration targets and calibrated parameter values are reported in the lower panel of Table 1.

It is important to notice that period 1 in GMO calibration does not refer to period 1 in our model since in our model people are 25 in period 1 whereas people are 20 in period 1 of GMO calibration. The period 1 in the calibration rather refers to the first period of calibrated self-control function (5). Thus, the β of the agent in the first period of his life (at age 25) is equal to β_6 according to (5), which equals 0.517. Similarly, we need to determine β for agents between the ages of 70 and 75 since the function (5) defines self-control problems only for the ages between 20 and 70. Given that according to Green *et al.* (1999) self-control problems already vanish by age 70, we set $\beta = 1$ for agents at ages between 70 and 75.

Now, we summarise our results.[16] First, in both of our calibrations, self-control problems decrease with age, and thus, it is always optimal to subsidise life-cycle savings. The red dots

[16]The `matlab` files used for all our simulations are available online. We refer to the README.pdf file for details.

and the solid green line on the left panel of Figure 2 represent the RR and GMO calibrations of the evolution of self-control problems over the life cycle, respectively. The corresponding lines on the right panel display the corresponding optimal subsidies. In the RR calibration, the optimal subsidies start about 6% and decrease slowly to about 1% at the end of the life cycle. In the GMO calibration, the subsidies start at a significantly higher 8.5% but decrease more sharply to 0% at the end of the life cycle. The optimality of declining subsidies in both cases is expected given Corollary 4 and Corollary 5 and the concavely increasing pattern of β_i with i in both calibrations. The fact that GMO optimal subsidies start higher and decline more steeply is due to the fact that the GMO self-control pattern is substantially more concave than that of RR.

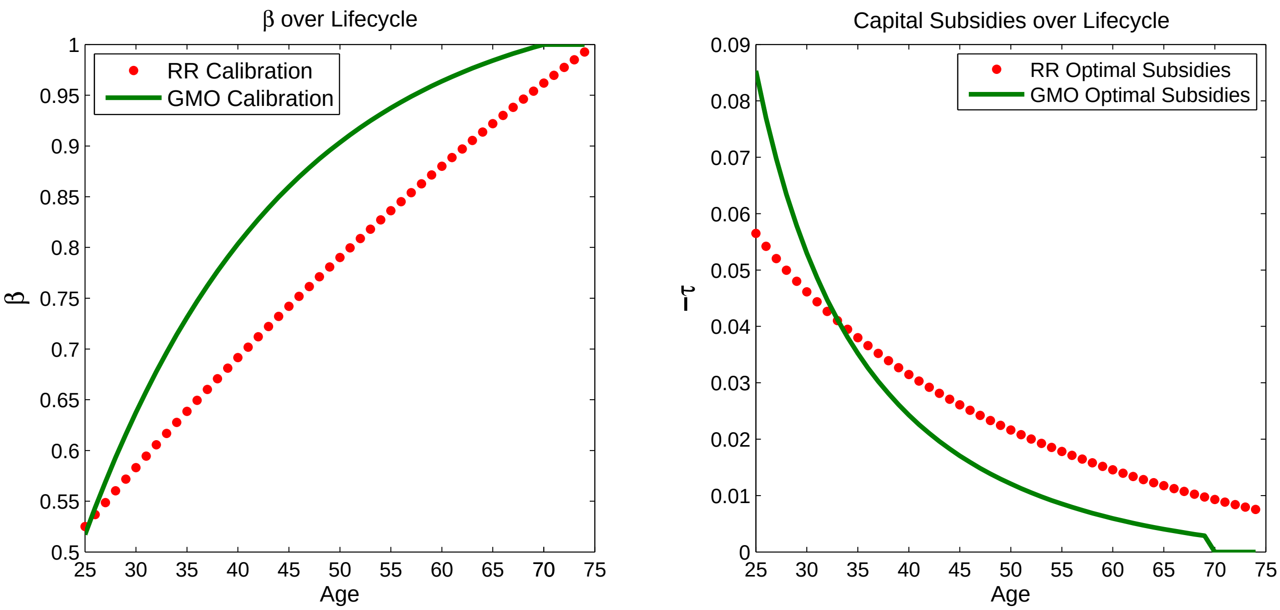


Figure 2: *Left panel: Life cycle patterns of self-control problems under Read and Read (2004) (RR) and Green* et al. *(1999) (GMO) calibrations. Right panel: Corresponding optimal subsidies.*

We conduct sensitivity analysis with respect to our choice of functional form for the self-control function given in (5). We do so by assuming a quadratic form: $\beta_i = Ai^2 + Bi + C$. For the RR calibration targets, the quadratic calibration gives a β_i pattern that is virtually identical to the original calibration. The implied optimal subsidies are virtually identical as well. For the GMO calibration targets, the implied β_i pattern rises strictly above one toward the end of the life cycle and stays above one until the end. This is inconsistent with the evidence found in both papers which calls for short term discount rates that are (weakly) larger than long term discount rates at all ages. The fact that quadratic function cannot be bounded above by one is one of the reasons why we have adopted the functional form in (5).[17]

We conclude this section by summarising the general pattern of optimal taxes that emerges in virtually all of our simulations: optimal taxes are negative, - i.e., they are in fact subsidies, and these subsidies typically decline with age.

2.3 How Large Are the Optimal Subsidies Relative to Existing Ones?

Observe that in our computations the tax base is the gross return on asset holdings. Most actual tax systems, however, tax *asset income*. If we translate our numbers into subsidies on capital income, we find that optimal subsidy at age 25 is 217%, and subsidies decline to 10% by age 65.[18] These are obviously large numbers. In this section, we compare these numbers with existing saving subsidies in the United States.

Tax-deferred retirement accounts are the main channel through which savings are subsidised in the United States. Probably, the most well-known tax-deferred saving account in the United States is the classical 401(k). Each dollar invested into a classical 401(k) is deductible from taxable income. Moreover, the taxes on the returns to 401(k) are deferred till retirement. This means, instead of paying taxes on interest or dividend income earned every year, the person pays tax based only on the income generated at the date of retirement, according to the tax rate faced by the agent at that date. Moreover, it is quite likely that, at retirement age, contributors face lower marginal taxes than when they invested into the plan. As we see below, this feature may generate considerable saving subsidies, and, importantly, these subsidies depend on where the agent is over her life cycle. Consider an agent who is at age i and is facing a marginal income tax rate τ_C^y based on the income tax bracket she falls in currently. Suppose there are N periods before she retires. If she invests \$1 today in 401(k), with the current tax deduction, this is as if she invests $\$\frac{1}{1-\tau_C^y}$. If τ_R^y is the income marginal tax rate at retirement age, the agent will receive

[17] See Pavoni and Yazici (2013a) for an extensive sensitivity analysis of optimal subsidies regarding the levels of β_1 and β_{avg} for the case of Green *et al.* (1999) calibration.

[18] Denoting the capital income tax by τ_i^k, the relation between our taxes and tax on capital income is given by: $1 - \tau_i^k = \frac{R(1-\tau_i)-1}{R-1}$. As a consequence, for $\tau_1 = -8.67\%$ and $\tau_{41} = -0.41\%$, and under $R = 1/1.04$, we have $\tau_1^k = -217\%$ and $\tau_{41}^k = -10\%$.

$\$R^N \frac{1-\tau_R^y}{1-\tau_C^y}$ at retirement. This implies a yearly saving subsidy of[19]

$$\left(\frac{1 - \tau_R^y}{1 - \tau_C^y} \right)^{\frac{1}{N}} - 1. \tag{6}$$

This saving subsidy depends on a person's age in two ways. First, people's income tend to depend on where they are on their life cycle, which implies the tax deduction they receive, τ_C^y, effectively depends on their age. Second, a person's age determines how far away she is from retirement, N, which clearly affects the subsidy rate in (6). Observe that in computing the implied 401(k) subsidy rate in (6), we take the tax base as the gross return on asset holdings to make it comparable to the optimal subsidies we have computed earlier in this section.

The 401(k) also allows employers to contribute to the worker retirement saving plans. The most common methods of employer matching are the \$1 per \$1 up to 6% of pay and the \$0.5 per \$1 up to 6% of pay.[20] According to these options, if the saving rate is below 6%, then for each dollar that a worker contributes to the 401(k) account, the employer contributes one (or, respectively, one half) dollar. This means a worker investing one dollar is effectively investing $\frac{\$2}{(1-\tau_C^y)}$ into the plan when the employer matching is 1-to-1 and $\frac{\$1.5}{(1-\tau_C^y)}$ when the employer matching is 0.5-to-1. The formula in (6) can then be straightforwardly adapted to compute implied 401(k) subsidies in the presence of employer matching.[21]

[19] The formula in (6) indicates that the implied subsidy rate of the 401(k) scheme is independent of the agent's saving rate. The 401(k) scheme, however, puts a cap on the amount agents can invest. In 2012, the maximal amount for agents aged 50 or below was \$17,000; older contributors faced an higher cap. The cap, however, is unlikely to be binding for the median household with an average saving rate.

[20] According to a 2009 survey conducted by Hewitt Associates, \$1 for \$1 up to 6% pay is the most common matching plan and is offered by 27% of all employers in their sample while \$0.5 per \$1 is the second most common matching plan. "Trends and Experience in 401(k) Plans." Retrieved from http://www.retirementmadesimpler.org/Library/Hewitt-Research-Trends-in-401k-Highlights.pdf

[21] Let a person's annual income be w and his amount contributed to 401(k) be x. If $x > 0.06w$, meaning the person is contributing more than 6% of his income, then in a one-to-one matching plan, the employer contributes $0.06w$ dollars, which implies that for each dollar he invests he is effectively investing

$$\frac{x + 0.06w}{x(1 - \tau_C^y)}.$$

Therefore, in this case, the implied subsidy depends on the amount contributed. However, Thaler and Benartzi (2004) report that the average saving rates into the SMarT plan (for the 'control group') are between 4.4% and 6.6% (see page S174). Thus, in our computations of the implied 401(k) subsidies, we assume that contribution rate is

Now, we compare the saving subsidies implied by a typical 401(k) plan to the optimal subsidies implied by our model. In Figure 3, right panel, we report the life-cycle profile of the median income per household head - between 25 and 66 years of age - in the period 2000-2006 and the corresponding marginal tax rates implied by the 2006 income tax code.[22] In Figure 3, left panel, we report the implied saving subsidies for several 401(k) plans together with the optimal subsidies given by our model under the two calibrations. Three observations are immediate. First, interestingly, the range of values for the subsidies implied by the 401(k) plan are not very different from the optimal ones. Second, the subsidies implied by the 401(k) plan are very much age-dependent. Third, the life-cycle pattern of the 401(k) subsidies is qualitatively very different from the optimal ones as they are *increasing over the life cycle*. Existing subsidies appear too low for young individuals and too high for individuals close to retirement.

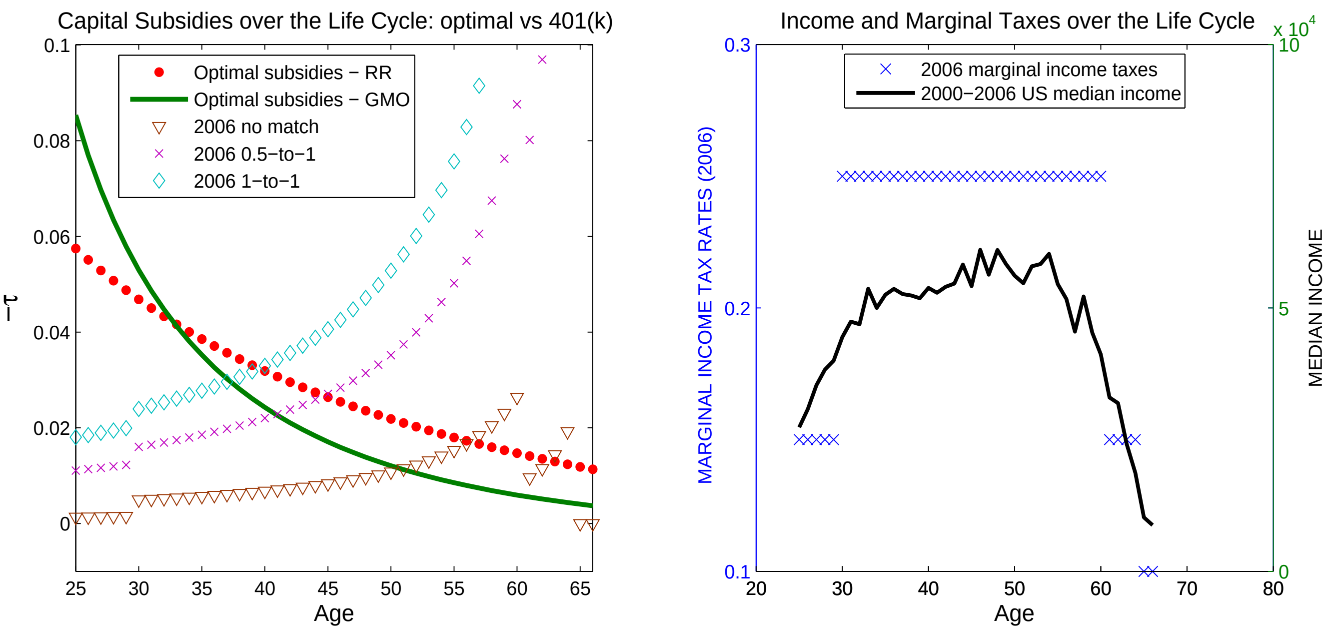


Figure 3: *Left panel: Optimal saving subsidies and subsidies implied by the 401(k) plan according to marginal income tax rates in 2006, at different levels of employer matching. Right panel: Median income of U.S. household head over life cycle in 2000-2006 and implied marginal income taxes in the year 2006.*

less than 6% and use the formula explained in the main text.

[22] The data for the life-cycle profile of the median income per household head in the period 2000-2006 is taken from Heathcote *et al.* (2010).

2.4 Comparison to the Constant Self-Control Model

In this section, we compare the optimal capital subsidies obtained in our model to those one would obtain in a model where self-control problems are constant over the life cycle. For the sake of brevity, we only discuss the comparison for the GMO calibration. The thick green line on the left panel of Figure 4 displays the GMO calibration whereas the thin (and flat) green line displays the corresponding constant self-control pattern in which people have $\beta = 0.818$ at all ages, which is the value that corresponds to average β in the GMO calibration.

The thin green line on the right-hand panel of Figure 4 displays that the optimal subsidies in the constant self-control model are constant at slightly below 1%, at 0.89% to be precise. In our GMO simulation, optimal capital subsidies start as high as 8.5% at the beginning of the life cycle. The fact that our subsidies are higher than those in the constant self-control model in the early years of the life cycle might not be very surprising since people have bigger self-control problems at earlier ages in our model. What is perhaps more surprising is that the level of optimal subsidies remain higher than the ones implied by the constant self-control model until as late as age 55 even though the left panel of Figure 4 shows that at 55 the agent in our model has significantly more self-control than the agent in the constant self-control model (with a β difference of about 0.1). This indicates that our model implies larger subsidies for similar levels of current self-control problems. The comparison at age 42 makes this point clearer: at 42, agents in both models have virtually the same level of $\beta \approx 0.82$. As the right panel of Figure 4 displays, the optimal subsidy in our model at this age, 2.14%, is much higher than the optimal subsidy in the constant self-control model, 0.89% (in terms of subsidies on capital income the comparison is 54% vs. 22%).

To see why the optimal subsidies in our model are significantly larger than the ones in constant self-control model, rewrite the optimal tax formula given by Proposition 1 as

$$-\tau_i^* = (1-\delta)\left(\frac{1}{\beta_i} - 1\right) + \delta \frac{\beta_{i+1} - \beta_i}{\beta_i}. \tag{7}$$

The expression (7) decomposes optimal subsidy formula into two components. The first component is the optimal subsidy that arises in a model if the self-control problem remains constant at the current level, β_i. The second component is the additional amount of subsidy needed due purely to the change in the level of self-control problems. Obviously, as long as self-control problem is decreasing with age, $\beta_{i+1} > \beta_i$, this term calls for additional subsidisation of savings. Since δ is typically close to one, the second component plays a quantitatively important role in shaping capital subsidies and taxes.

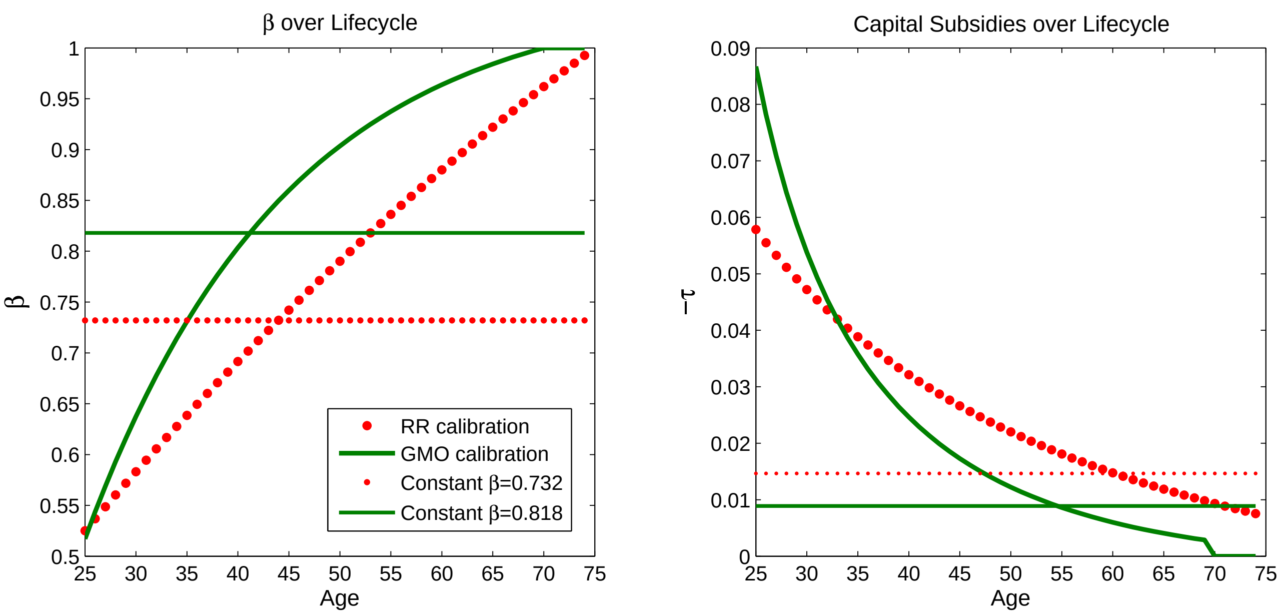


Figure 4: *Left panel: Life cycle patterns of self-control problems under Read and Read (2004) (RR) and Green* et al. *(1999) (GMO) calibrations, and under constant self-control model at* $\beta = 0.732$ *and* $\beta = 0.818$. *Right panel: Corresponding optimal subsidies.*

To grasp the intuition why our model implies higher subsidies, remember the decomposition of optimal taxes into the *current* and the *future* components, which we discuss following Proposition 2. The current component is related to an agent's current degree of self-control problem and calls for under-saving. For agent at age 42, this component is the same between our model and the constant self-control model. The future component summarises how much a person over-saves to compensate for future self's under-saving. Since people's degree of self-control improves with age in our model, the future component makes the agent save more today in the constant self-control model relative to ours. As a result, an agent with the same level of current self-control problem saves more in the constant self-control model, which implies the required subsidy to make him save the right amount is going to be lower. Due to this future component, optimal subsidies in our model remain higher than the one in the constant self-control model even long after age 42 (until age 55).

3 Imperfect Altruism and Life-Cycle Consumption Profile

In the main body of the paper, we have assumed that people are perfectly altruistic towards their descendants. In Section 3.1, we remove this assumption and investigate the quantitative importance of imperfect altruism on optimal subsidies. The main conclusion is that imperfect altruism has little effect on both the level and pattern of optimal saving subsidies. In Section 3.2, we extend the finite life time version of the model developed in Section 3.1 by allowing for borrowing constraints and use this model to study the life-cycle consumption implications of our model. We find that, even though our model abstracts from important life cycle issues such as child-rearing and health, life-cycle consumption profiles implied by our model under calibrated self-control patterns capture the key properties of empirical life-cycle consumption patterns fairly well.

3.1 Imperfect Altruism

Using a *constant self-control* model with sophisticated agents Krusell *et al.* (2010) find that optimal saving subsidies should be increasing with age if agents face finite life times. In our baseline model, the finite life time channel, which calls for increasing subsidies with age, is shut down by the perfect altruism assumption. We now consider an extended version of our model allowing for imperfect altruism and assess its quantitative importance in shaping optimal subsidies over the life cycle.

A parent has the following preferences over dynastic consumption streams

$$u(c_0) + \gamma\left[\delta u(c_1) + \delta^2 u(c_2) + \cdots + \delta^I u(c_I) + \delta^{I+1} u(c_0') + \delta^{I+1}\gamma\left[\delta u(c_1') + \delta^2 u(c_2') + \ldots\right]\ldots\right],$$

where this preference specification is equivalent to the one in the baseline model whenever the altruism factor, γ, is equal to 1. When $\gamma \in [0,1)$, there is imperfect altruism. The finite life time case of Krusell *et al.* (2010) corresponds to the case of $\gamma = 0$. The rest of the parent's problem is identical to the one in Section 1.3. Proposition 6 generalises the optimal tax formulas of Proposition 1 to the case with a general altruism factor, γ. In the case of perfect altruism, $\gamma = 1$, these formulas reduce back to the ones in Proposition 1.

Proposition 6 *Suppose $u(c) = log(c)$. For $\gamma \in (0,1]$, the optimal taxes are given by:*

$$1 - \tau_i^* = \frac{1}{\beta_i}\frac{1 + \beta_{i+1}\delta\left[1 + \delta + \cdots + \delta^{I-i-2} + \delta^{I-i-1}D\right]}{1 + \delta + \cdots + \delta^{I-i-1} + \delta^{I-i}D}, \textit{ for } i \in \{0, 1, \ldots, I\},$$

where $\beta_0 = \beta_{I+1} = 1$ *and*

$$D = \frac{1 + \delta\gamma(1 + \delta + \cdots + \delta^{I-1})}{1 - \delta^{I+1}\gamma}.$$

Proof. Relegated to Appendix A.3. ■

Figure 5 displays the quantitative effects of the finite life time channel on the monotonicity properties of optimal subsidies. The red dots represent optimal subsidies under RR calibration with $\gamma = 1$, whereas the dashed red line represents those under RR calibration but with $\gamma = 0$. The comparison of the two lines shows that the finite life time effect is not strong enough to overturn the optimality of subsidies declining with age. Similarly, the solid green line and the dashed green line in Figure 5 display optimal subsidies under GMO calibration with $\gamma = 1$ and $\gamma = 0$, respectively. The comparison in this case supports the earlier conclusion that optimal subsidies decline with age even when we take into account the finite life time effect. Observe that in both robustness exercises we set $\gamma = 0$. For $\gamma \in (0, 1)$, the effect of finite life time on the monotonicity properties of optimal subsidies would be even smaller.

We simulated our model adopting several different parameterisations of γ, δ and $i \to \beta_i$. The optimality of decreasing saving subsidies is quite robust (details are available upon request). What seems to play an important role to maintain robustness of the decreasing subsidy result under various values of γ is that, in our model, self-control problems vanish towards the end of the life cycle. Notice that the only way in which γ enters the subsidy formula in Proposition 6 is via D. Early in the life cycle, optimal subsidies are not sensitive to γ because the subsidy formula for early periods discounts D repeatedly. The optimal subsidy formula for later periods does not discount D heavily, but this time β_{i+1} converges to one, and the second fraction on the right-hand side of the subsidy formula converges to one, making optimal subsidies insensitive to γ.

Observe that, in Figure 5, for each calibration, the optimal subsidies are uniformly higher in the model with imperfect altruism than those implied by the model with perfect altruism. Mechanically, this is because the parameter γ enters into the formula for $1 - \tau_i^*$ in Proposition 6 only through the constant D. It is easy to show that D is increasing in γ, which implies that the subsidy is decreasing in γ since the subsidy is decreasing in D. Intuitively, the second term in the formula of $1 - \tau_i^*$ in Proposition 6 (the future component) is there because agent i disagrees with agent $i+1$ regarding how agent $i+1$ should discount consumption in period $i+2$ and onwards relative to consumption in $i+1$: from agent i's perspective, the correct discount factor between consumption at date $i+1$ and $i+s$ is δ^{s-1}, whereas agent $i+1$ discounts by $\beta_{i+1}\delta^{s-1}$. To correct for the eventual under-saving of self $i+1$, self i over-saves relative to the efficient allocation, and to prevent this, the government taxes self i. If $\gamma = 0$, the disagreement between self i and self $i+1$ regarding the discounting between $i+1$ consumption and future consumption levels stop at the end of the current life cycle ($D = 1$), while for $\gamma = 1$, the disagreement piles up for infinitely many generations ($D = \frac{1}{1-\delta}$). As a result, when $\gamma = 0$, there is less cumulative disagreement, which means self i is less motivated to over-save relative to the efficient allocation, which implies the tax implied by the future component is lower. Therefore, the overall subsidy is larger.

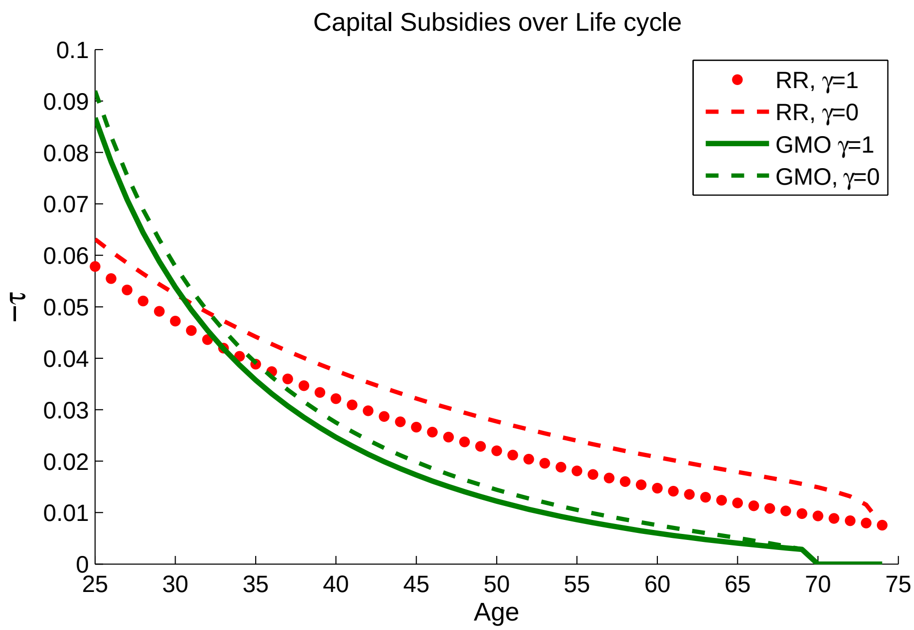


Figure 5: *Left panel: Life cycle patterns of self-control problems under Read and Read (2004) (RR) and Green* et al. *(1999) (GMO) calibrations with $\gamma = 0$ and $\gamma = 1$. Right panel: Corresponding optimal subsidies.*

3.2 Life-cycle Consumption Implications of Self-Control Calibration

Throughout the paper, we motivate and calibrate the life-cycle pattern of self-control using experimental data. In this section, we would like to document the life-cycle profile of consumption implied by our model as a partial validation of our calibration exercise.

We take Gourinchas and Parker (2002) (GP hereafter) as our empirical reference. GP computes life-cycle profiles of mean annual real disposable income and consumption using Consumer Expenditure Survey (CEX) data, employing synthetic cohort techniques.[23] They construct their variables from a sample of roughly 40,000 households from 1980 to 1993. A graphical representation of the income and consumption profiles they compute can be found in Figure 2 (page 67) in GP.

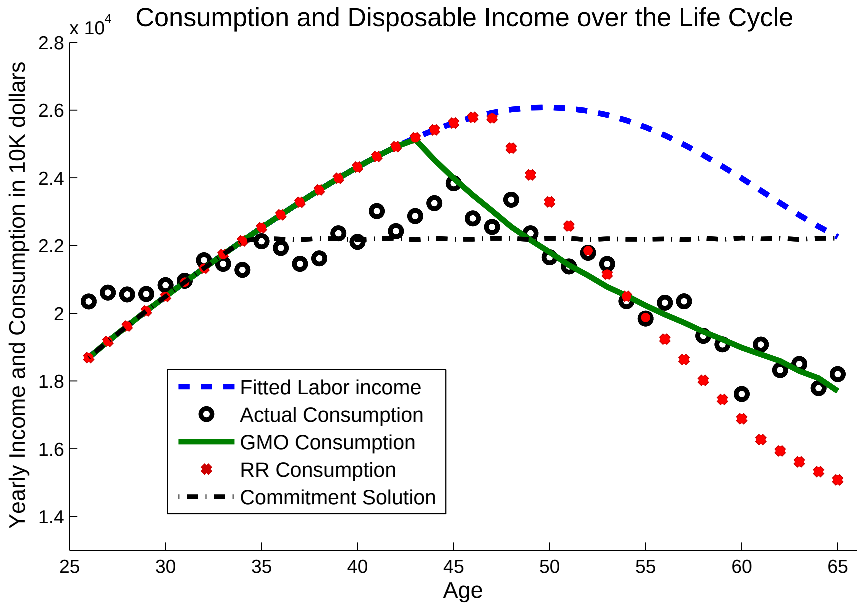


Figure 6: *The blue dashed line and the black circles represent the life-cycle profiles of mean disposable income and consumption between 1980 and 1993 as fitted by Gourinchas and Parker (2002). The red crosses and the green solid line represent the simulated consumption profiles implied by our model under self-control problems given by RR and GMO calibrations, respectively. The dash-dot line represents the life-cycle consumption profile of a planner who discounts exponentially and faces the same liquidity constraints that households face.*

Clearly, our simple model cannot have the ambition of matching the life-cycle pattern of consumption as closely as GP does. In particular, we abstract from income and health uncertainty, and from all life cycle changes other than income and self-control. We assume households face

[23]Consumption is obtained by subtracting expenditure on education, medical care, and mortgage interest payments from total household expenditure. Disposable labor income is obtained by subtracting Social Security tax payments, pension contributions, after tax asset and interest income, and those expenditures subtracted from consumption from after-tax family income.

a deterministic path of labor income which corresponds to the life-cycle profile of income constructed in GP. We also impose a constant (and exogenous) gross return to savings, R, and a simple form of liquidity constraints: households cannot borrow at all. To be consistent with GP, and given that our target is to get a life-cycle profile of consumption, we move away from the infinite horizon model and assume households are not altruistic ($\gamma = 0$). We use this model to compute two alternative life-cycle consumption profiles, one assuming self-control problems evolve according to our RR calibration and another one according to GMO calibration. We set the remaining parameters as follows. Agents are assumed to have logarithmic utility and we set $R = 1.0344$ and $\delta = \frac{1}{R}$.[24] Consistently with GP, agents start their working life at age 26 and retire at 65. People begin their lives with zero assets. After retirement, agents are assumed to leave for 10 more years and we set retirement income to zero.[25]

Figure 6 depicts life-cycle consumption profiles coming from RR and GMO simulations (red dots and solid green line, respectively), together with life-cycle (fitted) income and consumption profiles (blue dashed line and black circles, respectively) constructed in GP. All series are plotted for the age range between 26 and 65.[26] Figure 6 shows that both RR and GMO calibrations of self-control profiles imply reasonable life-cycle consumption profiles. In particular, our simple model of self-control is able to generate the concave pattern of consumption we observe in the data. It matches the declining pattern of consumption in the second part of the life cycle fairly well (for the GMO calibration the fit is, in fact, - surprisingly - quite remarkable), while

[24]We assume logarithmic utility mainly in order to use our closed form solutions; nonetheless, the relative risk aversion coefficient of one is well in the range estimated by GP, 0.5 to 1.4, depending on the exact specification of the model. The value of R is chosen to be consistent with GP, which estimates an average real return of 3.44% for Moody's AAA municipal bonds over the sample period.

[25] GP does not guide us regarding the length of retirement period or retirement income as it reports neither of them, and estimates a value of retirement and the implied 'retirement wealth' to best match moments on consumption and savings. Instead of choosing these values to optimise the model's fit, we set retirement income to zero and retirement length to 10 years. For both calibrations, the consumption profiles simulated using the model with 20 years of post retirement life - not report for ease of graphical exposition - are virtually identical to the profiles obtained from the 10 years model.

[26]Given the hump shaped pattern of income and the decreasing pattern of self-control, in all parametrisations, agents are liquidity constrained early in life. After they start saving, they never become liquidity constrained again. This characteristic of the model allows us to use the machinery we developed to solve the model without liquidity constraints (leading to the closed form solution for taxes) to solve this model with liquidity constraints. Further details on the numerical algorithm are available upon request.

it is unable to match the flatter pattern of consumption we observe in the data during the first part of the life cycle. GP argues that precautionary savings and measurement error play a crucial role in explaining the consumption profile in the initial segment of the life cycle while our model postulates deterministic income.[27]

It might be worthwhile discussing the optimal taxes implied by the model with liquidity constraints. The black dash-dot line in Figure 6 represents the life-cycle consumption profile of a planner who discounts exponentially and faces the same liquidity constraints that households face. This is the commitment consumption profile that optimal subsidies target under borrowing constraints. The figure shows that the consumption patterns of the commitment allocation and the self-control model (for both calibrations) coincide until age 34. The optimal subsidies in the borrowing constraint model starting with age 34 coincides with the optimal subsidies given by Proposition 6 (for $\gamma = 0$) since after this period the commitment allocation requires people to save. The optimal subsidies after age 34 corresponds to that shown by the thin dotted (green or red depending on the calibrations) lines in Figure 5. During the ages between 25 and 34, the optimal subsidies are indeterminate. Intuitively, as long as subsidies are not too high during these years, people will be borrowing constrained, and consume exactly their labor income, which is what we see in the commitment allocation. Notice that, the subsidies given by Proposition 6 are optimal for these years as well as they lie within the range of optimal subsidies.

4 Conclusion

This paper studies optimal capital taxation in an economy where agents face self-control problems. In line with evidence suggested by personality psychology and experimental studies, we allow for the severity of the self-control problem to change over the life cycle. We restrict attention to CIES utility functions and focus on linear Markov equilibria. We derive explicit formulas which allow us to compute optimal taxes given the evolution of self-control problem over the life cycle. We show that if agents ability to self-control increases concavely with age,

[27]On page 49, GP writes: *"The importance of the precautionary motive early in life implies that between 60 and 70 percent of non-pension wealth is due to precautionary savings, according to poor estimates and holding the real interest rate fixed."* On page 67, GP also writes: *"Consumption lies above income over the late twenties. Given that the CEX wealth data, and better household wealth surveys, show modest increases in liquid wealth over these ranges, this feature seems likely due to misreporting of income or consumption."*

then capital should be subsidised and the subsidy should decrease with age.

In our calibrated quantitative exercises, we find that optimal capital subsidies start somewhere between 6% and 8% at the beginning of the life cycle and decline monotonically with age to somewhere between 0% and 1%. If we translate the subsidies on capital into subsidies on capital income, these are very large numbers. Perhaps more importantly, we show they are much larger than the savings subsidy we would obtain in models with constant self-control, at most ages. Our model is probably too simple for delivering precise policy predictions. Nevertheless, our analysis suggests that researchers who take self-control problems seriously should also carefully measure the evolution of self-control problems over the life cycle before making policy suggestions.

We also compare our optimal subsidies with those implied by the 401(k) plan. If we exclude the very last periods before retirement - where the subsidy rate in the 401(k) essentially mimics the employer matching rate - the subsidy levels in the two cases are of comparable magnitudes. A marked difference emerges, however, in the life-cycle pattern of optimal subsidies: the 401(k) plan implies an increasing pattern of subsidies while the optimal subsidies decrease over the life cycle.

In Appendix C, we extend the model by allowing for different degrees of self awareness (partial sophistication) about the existence of future self-control problems. We first prove that when utility function is logarithmic, the optimal tax formulas are independent of the pattern of partial sophistication. When CEIS coefficient is different from one, closed form solutions for optimal taxes are unavailable. Our numerical experiments show that, as long as the level of sophistication is not changing abruptly from one period to another, the pattern of optimal capital subsidies over the life cycle is surprisingly robust to the degree of sophistication.

The existence of illiquid assets does not change our optimal tax results as long as there are no borrowing constraints. More precisely, in Appendix D we use a three periods example to show that a tax system that is optimal in an environment without illiquid assets is still optimal in the same environment with an illiquid asset as long as we complement the tax system with an appropriate tax on the illiquid asset.

Nicola Pavoni, Department of Economics, Bocconi University, via Roentgen 1, 20136 Milan, Italy

Hakki Yazici, Sabanci University, FASS, Orta Mahalle, Universite Caddesi No:27 Tuzla, 34956 Istanbul, Turkey

References

Ameriks, J., Caplin, A., Leahy, J. and Tyler, T. (2007). 'Measuring self-control problems', *American Economic Review*, vol. 97(3), pp. 966–972.

Ariely, D. and Wertenbroch, K. (2002). 'Procrastination, deadlines, and performance: Self-control by precommitment', *Psychological Science*, vol. 13(3), pp. 219–224.

Caspi, A., Roberts, B.W., Robins, R.W. and Trzesniewski, K.H. (2003). *Handbook of Life Course*.

DellaVigna, S. (2009). 'Psychology and economics: Evidence from the field', *Journal of Economic Literature*, vol. 47(2), pp. 315–372.

DellaVigna, S. and Malmendier, U. (2004). 'Contract design and self-control: Theory and evidence', *The Quarterly Journal of Economics*, vol. 119(2), pp. 353–402.

Erosa, A. and Gervais, M. (2002). 'Optimal taxation in life-cycle economies', *Journal of Economic Theory*, vol. 105(2), pp. 338–369.

Farhi, E. and Werning, I. (2013). 'Insurance and taxation over the life cycle', *Review of Economic Studies*, vol. 810(2), pp. 596–635.

Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002). 'Time discounting and time preference: A critical review', *Journal of Economic Literature*, vol. 40(2), pp. 351–401.

Golosov, M., Troshkin, M. and Tsyvinski, A. (2011). 'Optimal dynamic taxes', National Bureau of Economic Research, Inc.

Gourinchas, P.O. and Parker, J.A. (2002). 'Consumption Over the Life Cycle', *Econometrica*, vol. 70(1), pp. 47–89.

Green, L., Fry, A. and Myerson, J. (1994). 'Discounting of delayed rewards: A life-span comparison', *Psychological Science*, vol. 5, pp. 33–36.

Green, L., Myerson, J., Lichtman, D., Rosen, S. and Fry, A. (1996). 'Temporal discounting in choice between delayed rewards: The role of age and income', *Psychology and Aging*, vol. 11(1), pp. 79–84.

Green, L., Myerson, J. and Ostaszewski, P. (1999). 'Discounting of delayed rewards across the life-span: Age differences in individual discounting functions', *Behavioural Processes*, vol. 46(1), pp. 89–96.

Gruber, J. and Koszegi, B. (2004). 'Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes', *Journal of Public Economics*, vol. 88(9-10), pp. 1959–1987.

Gul, F. and Pesendorfer, W. (2001). 'Temptation and self-control', *Econometrica*, vol. 69(6), pp. 1403–1435.

Heathcote, J., Perri, F. and Violante, G.L. (2010). 'Unequal we stand: An empirical analysis of economic inequality in the united states: 1967-2006', *Review of Economic Dynamics*, vol. 13(1), pp. 15–51.

Helson, R., Jones, C. and Kwan, V.S.Y. (2002). 'Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples', *Journal of Personality and Social Psychology*, vol. 83(3), pp. 752–766.

Imrohoroglu, A., Imrohoroglu, S. and Joines, D.H. (2003). 'Time-inconsistent preferences and social security', *The Quarterly Journal of Economics*, vol. 118(2), pp. 745–784.

John, O.P., Gosling, S.D., Potter, J. and Srivastava, S. (2003). 'Development of personality in early and middle adulthood: Set like plaster or persistent change?', *Journal of Personality and Social Psychology*, vol. 84(5), pp. 1041–1053.

Krusell, P., Kuruscu, B. and Smith, A.J. (2002). 'Equilibrium welfare and government policy with quasi-geometric discounting', *Journal of Economic Theory*, vol. 105(1), pp. 42–72.

Krusell, P., Kuruscu, B. and Smith, A.J. (2010). 'Temptation and taxation', *Econometrica*, vol. 78(6), pp. 2063–2084.

Krusell, P. and Smith, A.A. (2003). 'Consumption–savings decisions with quasi–geometric discounting', *Econometrica*, vol. 71(1), pp. 365–375.

Laibson, D. (1994). 'Self-control and saving', *Ph.D. Dissertation*.

Laibson, D. (1997). 'Golden eggs and hyperbolic discounting', *The Quarterly Journal of Economics*, vol. 112(2), pp. 443–477.

Laibson, D., Repetto, A. and Tobacman, J. (2007). 'Estimating discount functions with consumption choices over the lifecycle', National Bureau of Economic Research.

O'Donoghue, T. and Rabin, M. (2003). 'Studying optimal paternalism, illustrated by a model of sin taxes', *American Economic Review*, vol. 93(2), pp. 186–191.

O'Donoghue, T. and Rabin, M. (2006). 'Optimal sin taxes', *Journal of Public Economics*, vol. 90(10-11), pp. 1825–1849.

Pavoni, N. and Yazici, H. (2013a). 'Optimal life-cycle capital taxation under self-control problems', .

Pavoni, N. and Yazici, H. (2013b). 'Present-bias preferences and optimal taxation of parental transfers', .

Phelps, E.S. and Pollak, R.A. (1968). 'On second-best national saving and game-equilibrium growth', *The Review of Economic Studies*, vol. 35(2), pp. 185–199, ISSN 00346527.

Read, D. and Read, N. (2004). 'Time discounting over the lifespan', *Organizational Behavior and Human Decision Processes*, vol. 94(1), pp. 22–32.

Roberts, B.W., Walton, K.E. and Viechtbauer, W. (2006). 'Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies', *Psychological Bulletin*, vol. 132(1), pp. 1–25.

Strotz, R.H. (1955). 'Myopia and inconsistency in dynamic utility maximization', *The Review of Economic Studies*, vol. 23(3), pp. 165–180.

Thaler, R.H. and Benartzi, S. (2004). 'Save more tomorrow (tm): Using behavioral economics to increase employee saving', *Journal of Political Economy*, vol. 112(S1), pp. S164–S187.

Thaler, R.H. and Shefrin, H.M. (1981). 'An economic theory of self-control', *Journal of Political Economy*, vol. 89(2), pp. 392–406.

Optimal Life-cycle Capital Taxation under Self-Control Problems

ONLINE APPENDIX

Nicola Pavoni* and Hakki Yazici†

August 11, 2015

A Proofs

A.1 Proof of Proposition 1.

In this section, we provide the proof of our main result, Proposition 1, for the general setup where the economy starts from any initial level of capital stock and prices change over time. In order to do so, we first define the parent's problem under taxes in the general setup.

Preparation to the proof.

Let k_0 be the initial level of capital stock and $\{k_t^*\}_t$ be the sequence of the efficient capital levels that start from k_0. We know that the commitment allocation is recursive in k_t. Let $K : \mathbb{R} \to \mathbb{R}$ be the function describing the evolution of the aggregate level of capital in the commitment allocation:

$$k_{t+1}^* = K(k_t^*).$$

Agents face a price sequence satisfying:

$$\begin{aligned} R(k_t) &= f'(k_t), \\ w(k_t) &= f(k_t) - f'(k_t)k_t, \end{aligned}$$

that is, it is generated by a capital stock sequence $\{k_t^*\}_t$ where the capital stock is generated by K. Since the problem is recursive, a government which aims to implement the efficient allocation will use the same taxes in any two periods if the age of the agent and the capital stock in those periods are the same. Therefore, without loss of generality, we define taxes as functions of age and capital stock as

*Bocconi University, IGIER, IFS, and CEPR. E-mail: nicola.pavoni@unibocconi.it

†Sabanci University. E-mail: hakkiyazici@sabanciuniv.edu

follows: $\tau_i(k_t)$ is the savings (capital) tax agent at age $i = 0, 1, .., I$ pays if the capital stock in that period is k_t. Government (per-period) budget feasibility requires the lump-sum rebate to satisfy: $T_i(k_t) = R(k_t)\tau_i(k_t)b_i(k_t;\tau)$.

To describe the problem of the agents, we define the policy functions $b_i(\cdot, k_t; \tau)$ describing the optimal behaviour of the agent i as function of b_{i-1} given the level of aggregate capital k_t, the taxes $\tau := \{\tau_i(\cdot), T_i(\cdot)\}_i$ and what he believes other agents' rules will be, and that the evolution of capital follows the rule K. When agent n is deciding b_n, his evaluation of the effect of his choice on b_i, $i > n$ will be described by the function $b_i(b_{i-1}(...b_{n+1}(b_n, k_t^*; \tau)...), k_{t+i-n-1}^*; \tau), k_{t+i-n}^*; \tau)$, where for all t, s, we define $k_{t+s}^* = K(K(...(k_t^*)...))$, where the K function has been applied s times. To simplify notation, we will denote this mapping simply as $b_i(..(b_n)..))$.

Finally, our notation will be simplified if we let k be the level of capital stock already in place in the last period of a parent and k' or k^1 refer to the capital stock next period and k^i refer to the level of capital stock i periods after the period in which capital stock was k, namely: $k^i = K(K(...(k)...))$, where the function K has been applied i times. In the problem below, the function K is fixed to that of the commitment allocation. Of course, the function describing the evolution of aggregate capital in equilibrium is part of the fixed point argument as it must satisfy market clearing.

Parent's Problem along the Transition

$$
\begin{aligned}
V(b,k;\tau) \quad &= \quad \max_{b_0} \; u\left(R(k)\left(1-\tau_I\right)b + w(k) + T_I - b_0\right) \\
&+\delta\left[\sum_{i=0}^{I-1} \delta^i u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(b_0)...) + w(k^{i+1}) + T_i - b_{i+1}\right) + \delta^I V\left(b_I(...(b_0)...), k^{I+1}; \tau\right)\right] \\
&\text{s.t. for all } \; b_0
\end{aligned}
$$

$$
b_1(b_0, k^1; \tau) = \arg\max_{\hat{b}_1} \; u\left(R(k^1)\left(1-\tau_0\right)b_0 + w(k^1) + T_0 - \hat{b}_1\right)
$$

$$
+\delta\beta_1\left[\; \textstyle\sum_{i=1}^{I-1} \; \delta^{i-1} u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(\hat{b}_1)...) + w(k^{i+1}) + T_i - b_{i+1}(...(\hat{b}_1)...)\right) + \delta^{I-1} V\left(b_I(...(\hat{b}_1)...), k^{I+1}; \tau\right) \;\right]
$$

$$
\text{s.t. for all } \; b_1
$$

$$
b_2(b_1, k^2; \tau) = \arg\max_{\hat{b}_2} \; u\left(R(k^2)\left(1-\tau_1\right)b_1 + w(k^2) + T_1 - \hat{b}_2\right)
$$

$$
+\delta\beta_2\left[\; \textstyle\sum_{i=2}^{I-1} \; \delta^{i-2} u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(\hat{b}_2)...) + w(k^{i+1}) + T_i - b_{i+1}(...(\hat{b}_2)...)\right) + \delta^{I-2} V\left(b_I(...(\hat{b}_2)...), k^{I+1}; \tau\right) \;\right]
$$

$$
\text{s.t. for all } \; b_2
$$

$$
...
$$

$$
\text{s.t. for all } \; b_{I-2}
$$

$$
b_{I-1}(b_{I-2}, k^{I-1}; \tau) \in \arg\max_{\hat{b}_{I-1}} \; u\left(R(k^{I-1})\left(1-\tau_{I-2}\right)b_{I-2} + w(k^{I-1}) + T_{I-2} - \hat{b}_{I-1}\right)
$$

$$
+\delta\beta_{I-1}\left[u\left(R(k^I)\left(1-\tau_{I-1}\right)\hat{b}_{I-1} + w(k^I) + T_{I-1} - b_I(...(\hat{b}_{I-1})...)\right) + \delta V\left(b_I(...(\hat{b}_{I-1})...), k^{I+1}; \tau\right)\right]
$$

$$
\text{s.t. for all } \; b_{I-1}
$$

$$
b_I(b_{I-1}, k^I; \tau) = \arg\max_{\hat{b}_I} \; u\left(R\left(1-\tau_{I-1}\right)b_{I-1} + w_{I-1} + T_{I-1} - \hat{b}_I\right) + \delta\beta_I V\left(\hat{b}_I, k^{I+1}; \tau\right),
$$

where $b_i(...(b_i)...) = b_i$.

Letting b_i and k^{i+1} be the saving level in period i and aggregate capital stock in period $i+1$, define (we disregard the tax dependence for notational simplicity):

$$\begin{aligned}
\Gamma_i(b_i, k^{i+1}) &= R(k^{i+1})(1-\tau_i(k^{i+1}))b_i + w(k^{i+1}) + T_i(k^{i+1}) + G_i(k^{i+1}), \\
G_i(k^{i+1}) &= \frac{T_{i+1}(k^{i+2}) + w(k^{i+2})}{R(k^{i+2})\left(1-\tau_{i+1}(k^{i+2})\right)} + \frac{T_{i+2}(k^{i+3}) + w(k^{i+3})}{\prod_{j=i+2}^{i+3} R(k^j)\left(1-\tau_{j-1}(k^j)\right)} + ... + \frac{T_I(k^{I+1}) + w(k^{I+1})}{\prod_{j=i+2}^{I} R(k^j)\left(1-\tau_{j-1}(k^j)\right)} + ..., \\
c_{i+1}(b_i, k^{i+1}) &= M_{i+1}\Gamma_i(b_i, k^{i+1}),
\end{aligned}$$

where $G_i(k^{i+1})$ is the total net present value of future lump-sum taxes and wages, and $\Gamma_i(b_i, k^{i+1})$ is the net present value of wealth available to agent at the beginning of age $i+1$ when the level of aggregate capital stock today is k^{i+1}, the agent saved b_i in the previous period, and M_{i+1} is the fraction consumed out of that wealth. It follows from the flow budget constraint in period $i+1$ that if the stated consumption rule is part of an optimal policy, agent's saving in period $i+1$ must satisfy for all b_i:

$$b_{i+1}(b_i, k^{i+1}; \tau) = R(k^{i+1})\left(1-\tau_i(k^{i+1})\right)b_i + w(k^{i+1}) + T_i(k^{i+1}) - M_{i+1}\Gamma_i(b_i, k^{i+1}).$$

Note that, using

$$\frac{\partial b_{i+1}(b_i, k^{i+1}; \tau)}{\partial b_i} = R(k^{i+1})\left(1-\tau_i(k^{i+1})\right) - M_{i+1}\frac{\partial \Gamma_i(b_i, k^{i+1})}{\partial b_i} = (1-M_{i+1})R(k^{i+1})\left(1-\tau_i(k^{i+1})\right),$$

it is relatively simple algebra to show that, under the consumption rule given above, net present value of wealth between any two consecutive periods is related as follows: for all $i = 1, ..., I$

$$\begin{aligned}
\Gamma_i(b_i(b_{i-1}, k^i; \tau), k^{i+1}) &= R(k^{i+1})(1-\tau_i(k^{i+1}))(1-M_i)\Gamma_{i-1}(b_{i-1}, k^i) \\
&\quad \textit{and} \\
\Gamma_0(b_0(b, k; \tau), k^1) &= R(k^1)(1-\tau_0(k^1))(1-M_0)\Gamma_I(b, k),
\end{aligned} \tag{1}$$

where

$$\Gamma_I(b, k) = R(k)(1-\tau_I(k))b + w(k) + T_I(k) + G_I(k)$$

is the net present value of wealth available to the parent when the level of aggregate capital stock today is k and the parent saved b in the previous period.

Using the above recursion, it is possible to express consumption as follows:

$$c_{i+1}(b_i(..(b)..), k^{i+1}) = Q_i(k)M_{i+1}\Gamma_I(b, k),$$

where $b_i(...(b)...)$ is the shortcut for the nested policy we describe above and

$$Q_i(k) := \Pi_{s=0}^{i}(1-M_s)\, R(k^{s+1})\left(1-\tau_s(k^{s+1})\right),$$

with $k^{s+1} = K(...(k)...)$, where the map K is applied $s+1$ times as usual.

Now using linearity of the policy functions and the first-order approach, we can rewrite the parent's problem as:

$$V\left(b,k;\tau\right) = \max_{M_0}\; u\left(M_0\Gamma_I\left(b\right)\right) + \delta\left[\sum_{i=0}^{I-1}\delta^i u\left(Q_i(k)M_{i+1}\Gamma_I\left(b\right)\right) + \delta^I V\left((1-M_I)Q_{I-1}(k)\Gamma_I\left(b\right), k^{I+1};\tau\right)\right] \quad (2)$$

$$s.t. \quad \text{for all } i \in \{1,...,I-1\}$$

$$\left(M_iQ_{i-1}(k)\Gamma_I\left(b,k\right)\right)^{-\sigma} = \delta\beta_i\left[R(k^{i+1})(1-\tau_i(k^{i+1}))\left\{\begin{array}{c}\sum_{j=i+1}^{I}\delta^{j-(i+1)}\left(M_jQ_{j-1}(k)\Gamma_I\left(b,k\right)\right)^{-\sigma}M_j\frac{Q_{j-1}(k)}{Q_i(k)}\\ +\delta^{I-i}V'(b_I(..(b)..),k^{I+1};\tau)(1-M_I)\frac{Q_{I-1}(k)}{Q_i(k)}\end{array}\right\}\right]$$

$$\left(M_IQ_{I-1}(k)\Gamma_I\left(b,k\right)\right)^{-\sigma} = \delta\beta_I V'\left(b_I(..(b)...),k^{I+1};\tau\right).$$

Core proof of Proposition 1.

We will prove that facing the sequence of efficient capital levels and the taxes specified in Proposition 1, people will choose the efficient allocation, thereby verifying both (1) that the sequence of the efficient capital levels is actually part of equilibrium under the taxes described in Proposition 1, and (2) that under the taxes specified by Proposition 1, people choose the efficient allocation.

Guess

$$V\left(b,k;\tau\right) = D\log(\Gamma_I\left(b,k\right)) + B(k),$$

where D and B are constants of the parent's value function.

Now, we compute the coefficients for parent's value function, D.

Compute V' in terms of D using the guess for value function above:

$$V'(b_I(..(b)..),k^{I+1};\tau) = DR(k^{I+1})(1-\tau_I(k^{I+1}))(\Gamma_I(b,k)Q_I(k))^{-1}, \quad (3)$$

where we used the recursion (1).

Plugging (3) in the constraints described in problem (2) and using the definition of Q_i, these constraints become: for all $i \in \{1,...,I-1\}$:

$$\left(M_iQ_{i-1}(k)\right)^{-1} = \delta\beta_iR(k^{i+1})(1-\tau_i(k^{i+1}))\left(Q_i(k)\right)^{-1}\left[\sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\right],$$

$$\textit{and}$$

$$\left(M_IQ_{I-1}(k)\right)^{-1} = \delta\beta_IR(1-\tau_I(k^{I+1}))\left(Q_I(k)\right)^{-1}D.$$

Now, using the marginal condition describing self-I behaviour, it is easy to show that

$$M_I(D) = \frac{1}{1+\beta_I\delta D}.$$

Similarly, use other constraints defining the policies to compute $M_i(D)$ for $i = 1,..,I-1$:

$$M_i(D) = \frac{1}{1+\beta_i\delta\left(\sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\right)}.$$

Taking first-order condition with respect to bequests in the parent's problem (2) and plugging in the $M_i(D)$ from above, we get:

$$M_0(D) = \frac{1}{1+\delta\left(\sum_{j=0}^{I-1}\delta^j + \delta^I D\right)}.$$

Now, we verify the value function to compute D :

$$\begin{aligned} D\log\left(\Gamma_I\left(b,k\right)\right) + B(k) &= \log\left(M_0(D)\Gamma_I\left(b,k\right)\right) \\ &+\delta\left[\sum_{i=0}^{I-1}\delta^i \log\left(Q_i(k)M_{i+1}(D)\Gamma_I\left(b,k\right)\right) + \delta^I\left\{D\log\left(\Gamma_I\left(b,k\right)Q_I(k)\right) + B(k^{I+1})\right\}\right], \end{aligned}$$

which implies

$$D = \sum_{i=0}^{I}\delta^i + \delta^{I+1}D$$

and hence

$$D = \frac{1}{1-\delta}.$$

By plugging D in the formula for $M_i(D)$, we compute

$$\begin{gathered} M_i = \frac{1-\delta}{1-\delta+\beta_i\delta}, \text{ for all } i \in \{1,..,I\}, \\ M_0 = 1-\delta. \end{gathered} \tag{4}$$

Now we turn to taxes that implement the efficient allocation. The constraint that describes self-i's behaviour for $i \in \{1,..,I-1\}$ becomes the following once we plug in the derivative of the value function from (3) :

$$\left(M_iQ_{i-1}(k)\Gamma_I\left(b,k\right)\right)^{-1} = \delta\beta_iR(k^{i+1})(1-\tau_i(k^{i+1}))\left(M_{i+1}Q_i(k)\Gamma_I\left(b,k\right)\right)^{-1}\left[\ \sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\ \right]M_{i+1}. \tag{5}$$

The comparison of (5) with the efficiency condition (1) in the main text gives the optimal tax as:

$$\begin{aligned} \left(1-\tau_i^*(k^{i+1})\right) &= \frac{1}{\beta_i}\left(\left[\ \sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\ \right]M_{i+1}\right)^{-1} \\ &= \frac{1}{\beta_i}\left(1-\delta+\beta_{i+1}\delta\right). \end{aligned}$$

For self-I, the constraint describing his behavior in problem (2) reads as follows:

$$\left(M_IQ_{I-1}(k)\Gamma_I\left(b,k\right)\right)^{-1} = \delta\beta_IR(k^{I+1})(1-\tau_I(k^{I+1}))\left(M_0Q_I(k)\Gamma_I\left(b,k\right)\right)^{-1}DM_0,$$

and the comparison of this with the efficiency condition gives

$$\left(1-\tau_I^*(k^{I+1})\right) = \frac{1}{\beta_I}.$$

Finally, a comparison of the following first-order condition of the parent

$$(M_0\Gamma_I(b,k))^{-1} = \delta R(k^1)(1-\tau_0(k^1))(M_1 Q_0(k)\Gamma_I(b,k))^{-1}\frac{\left[\sum_{i=0}^{I-1}\delta^i + \delta^I D\right]}{M_1^{-1}}$$

with the corresponding optimality condition gives

$$1-\tau_0^*(k^1) = (1-\delta+\beta_1\delta).$$

A.2 Proof of Proposition 2.

If we plug in the constraint defining the policy of the agent at age $i+1$ in the constraint of agent at age i, we get:

$$u'(c_i) = \delta\beta_i R(1-\tau_i)u'(c_{i+1})\left\{1+\frac{\partial b_{i+1}(b_i)}{\partial b_i}\frac{\left(\frac{1}{\beta_{i+1}}-1\right)}{R(1-\tau_i)}\right\},$$

which renders optimal taxes as:

$$(1-\tau_i^*) = \frac{1}{\beta_i}\frac{1}{1+\frac{\partial b_{i+1}(b_i^*)}{\partial b_i}\frac{\left(\frac{1}{\beta_{i+1}}-1\right)}{R(1-\tau_i^*)}}.$$

Under CEIS utility and linear policies, we have:

$$\frac{\partial b_{i+1}(b_i)}{\partial b_i} = (1-M_{i+1})R(1-\tau_i).$$

Now plug this in the tax formula above to get the CEIS specific tax formula:

$$(1-\tau_i^*) = \frac{1}{\beta_i}\frac{1}{1+(1-M_{i+1}^*)\left(\frac{1}{\beta_{i+1}}-1\right)}. \tag{6}$$

When $R\delta = 1$, in the efficient allocation we have $c_i^* = c_{i+1}^*$ for all i. This means

$$c_i^* = M_i^*\Gamma_{i-1}(b_{i-1}^*) = c_{i+1}^* = M_{i+1}^*\Gamma_i(b_i^*)$$

which, using the relationship $\Gamma_i(b_i) = R(1-\tau_i)(1-M_i)\Gamma_{i-1}(b_{i-1})$ implies

$$M_i^* = \frac{M_{i+1}^* R(1-\tau_i^*)}{1+M_{i+1}^* R(1-\tau_i^*)}. \tag{7}$$

Plugging (6) in (7), we get a system of $(I+1)$ equations in $(I+1)$ unknowns $(M_0^*, ..., M_I^*)$ that fully pin down agents policies when they face optimal taxes, for the CEIS case:

$$M_i^* = \frac{M_{i+1}^* R\frac{1}{\beta_i}\frac{1}{1+(1-M_{i+1}^*)\left(\frac{1}{\beta_{i+1}}-1\right)}}{1+M_{i+1}^* R\frac{1}{\beta_i}\frac{1}{1+(1-M_{i+1}^*)\left(\frac{1}{\beta_{i+1}}-1\right)}}$$

Clearly, the solution to this system does not depend on σ. In fact, it is easy to show that the logarithmic utility solution given by equation (4) satisfies the above system of equations, meaning it is an equilibrium. Plugging (4) in the formula for taxes, (6), we get that optimal taxes are the same as the logarithmic utility case.

A.3 Proof of Proposition 6.

The proof of Proposition 6 follows the proof of Proposition 1 very closely. The important difference is that the altruism factor, γ, can be any number in $[0,1]$. In this case, the maximiation problem of the parent is identical to (2), except that the objective function has the general altruism factor:

$$V\left(b,k;\tau\right)=\max_{M_0}\ u\left(M_0\Gamma_I\left(b\right)\right)+\gamma\delta\left[\sum_{i=0}^{I-1}\delta^i u\left(Q_i(k)M_{i+1}\Gamma_I\left(b\right)\right)+\delta^I V\left((1-M_I)Q_{I-1}(k)\Gamma_I\left(b\right),k^{I+1};\tau\right)\right]. \tag{8}$$

We will prove that facing the sequence of efficient capital levels and the taxes specified in Proposition 6, people will choose the efficient allocation, thereby verifying both (i) that the sequence of the efficient capital levels is actually part of equilibrium under the taxes described in Proposition 6, and (ii) that under the taxes specified by Proposition 6, people choose the efficient allocation.

Guess

$$V\left(b,k;\tau\right)=D\log(\Gamma_I\left(b,k\right))+B(k),$$

where D is the constant of the parent's value function.

Compute V' in terms of D using the guess for value function:

$$V'(b_I(..(b)..),k^{I+1};\tau)=DR(k^{I+1})(1-\tau_I(k^{I+1}))(\Gamma_I(b,k)Q_I(k))^{-1},$$

where we used the recursion (1).

Plugging these in the constraints described in problem (2), we get for all $i\in\{1,...,I-1\}$:

$$\begin{aligned}(M_iQ_{i-1}(k))^{-1} &= \delta\beta_i R(k^{i+1})(1-\tau_i(k^{i+1}))\left(Q_i(k)\right)^{-1}\left[\sum_{j=i+1}^{I}\delta^{j-(i+1)}+\delta^{I-i}D\right]\\ &\quad\ \textit{and}\\ (M_IQ_{I-1}(k))^{-1} &= \delta\beta_I R(1-\tau_I(k^{I+1}))\left(Q_I(k)\right)^{-1}D.\end{aligned}$$

Now, using the marginal condition describing self-I behaviour, it is easy to show that

$$M_I(D)=\frac{1}{1+\beta_I\delta D}.$$

Similarly, use other constraints defining the policies to compute $M_i(D)$ for $i=1,..,I-1:$

$$M_i(D)=\frac{1}{1+\beta_i\delta\left\{\sum_{j=i+1}^{I}\delta^{j-(i+1)}+\delta^{I-i}D\right\}}. \tag{9}$$

Taking first-order condition with respect to bequests in the parent's problem (2) and plugging in the $M_i(D)$ from above for all i, we get:

$$M_0(D)=\frac{1}{1+\delta\left(\sum_{j=0}^{I-1}\delta^j+\delta^I D\right)}.$$

Now verify the value function to compute D :

$$\begin{aligned} D\log\left(\Gamma_I\left(b,k\right)\right)+B(k) &= \log\left(M_0(D)\Gamma_I\left(b,k\right)\right) \\ &+\gamma\delta\left[\sum_{i=0}^{I-1}\delta^i\log\left(Q_i(k)M_{i+1}(D)\Gamma_I\left(b,k\right)\right)+\delta^I\left\{D\log\left(\Gamma_I\left(b,k\right)Q_I(k)\right)+B(k^{I+1})\right\}\right], \end{aligned}$$

which implies

$$D=1+\gamma\delta\left(\sum_{i=0}^{I-1}\delta^i+\delta^I D\right)$$

and hence

$$D=\frac{1+\gamma\delta\sum_{i=0}^{I-1}\delta^i}{1-\delta^{I+1}\gamma}.$$

Now we turn to taxes that implement the efficient allocation. The constraint that describes self-i's behaviour for $i\in\{1,..,I-1\}$ becomes the following once we plug in the derivatives of the value functions from (3) :

$$\left(M_iQ_{i-1}(k)\Gamma_I\left(b,k\right)\right)^{-1}=\delta\beta_iR(k^{i+1})(1-\tau_i(k^{i+1}))\left(M_{i+1}Q_i(k)\Gamma_I\left(b,k\right)\right)^{-1}\left[\sum_{j=i+1}^{I}\delta^{j-(i+1)}+\delta^{I-i}D\right]M_{i+1}. \tag{10}$$

The comparison of (10) with the efficiency condition (1) in the main text gives the optimal tax as:

$$1-\tau_i^*(k^{i+1})=\frac{1}{\beta_i}\left(\left[\sum_{j=i+1}^{I}\delta^{j-(i+1)}+\delta^{I-i}D\right]M_{i+1}\right)^{-1},$$

which, using (9), implies

$$1-\tau_i^*(k^{i+1})=\frac{1}{\beta_i}\frac{1+\beta_{i+1}\delta\left(1+\delta+...+\delta^{I-i-2}+\delta^{I-i-1}D\right)}{1+\delta+...+\delta^{I-i-1}+\delta^{I-i}D}.$$

For self-I, the constraint describing his behaviour in problem (2) reads as follows:

$$\left(M_IQ_{I-1}(k)\Gamma_I\left(b,k\right)\right)^{-1}=\delta\beta_IR(k^{I+1})(1-\tau_I(k^{I+1}))\left(M_0Q_I(k)\Gamma_I\left(b,k\right)\right)^{-1}DM_0,$$

and the comparison of this with the efficiency condition gives

$$1-\tau_I^*(k^{I+1})=\frac{1}{\beta_I}.$$

Finally, a comparison of the following first-order condition of the parent

$$\left(M_0\Gamma_I\left(b,k\right)\right)^{-1}=\gamma\delta R(k^1)(1-\tau_0(k^1))(M_1Q_0(k)\Gamma_I\left(b,k\right))^{-1}\frac{\left[\sum_{i=0}^{I-1}\delta^i+\delta^I D\right]}{M_1^{-1}}$$

with the corresponding optimality condition gives

$$1-\tau_0^*(k^1)=\frac{1+\beta_1\delta\left(1+\delta+...+\delta^{I-2}+\delta^{I-1}D\right)}{1+\delta+...+\delta^{I-1}+\delta^I D}.$$

B Approximating Hyperbolic Discount Functions with Quasi-hyperbolic Discount Functions

Green *et al.* (1999) and Read and Read (2004) are two studies that collect experimental data and use it to estimate inter-temporal discount functions for different age groups. In this section, we explain how we approximate our quasi-hyperbolic discount functions for those age groups using the hyperbolic discount functions estimated in Green *et al.* (1999) and Read and Read (2004).

Green *et al.* (1999) estimates (11) for two adult age groups (young and old adults). Read and Read (2004) estimates (11) for three adult age groups (young, middle-aged, and old). A key finding in both Green *et al.* (1999) and Read and Read (2004) is that the old adults groups in both studies discount future exponentially.

For the rest of the age groups, both papers find that the following class of hyperbola-like functions provide the best description for how each group discounts delayed rewards:

$$\zeta(D) = \frac{1}{(1 + kD)^s}, \tag{11}$$

where D is the length of delay to a future reward (measured in years) and k and s are the parameters that govern the rate of discounting and the scaling of amount and or delay. We take the hyperbolic discount function estimated for each age group and find the best approximation to that function within the set of quasi-hyperbolic discount functions that are parameterised by two parameters, δ and β. As we do all throughout the paper, we follow Laibson *et al.* (2007) and set $\delta = 0.96$.

To see how we approximate $\beta's$, let us focus on the young adult group in Read and Read (2004) as an example. Read and Read (2004) estimate $k = 0.076$ and $s = 0.516$ for this age group. We first simulate yearly discount factors as a function of years of delay implied by the hyperbolic discount function estimated for this age group. Then, we set $\delta = 0.96$ choose β using a simple least squares procedure: that is, we choose β to minimise the sum of the squares of errors between the yearly discount factors that are implied by the hyperbolic discount function and the quasi-hyperbolic discount function. For the young adult group in Read and Read (2004), this procedure gives us $\beta = 0.525$. We repeat this procedure for each age group in each study. The table below summarises the approximation procedure.

Table 1: **Approximating β from Hyperbolic Discount Functions**

Age group	k	s	β
Young adults (RR)	0.076	0.516	0.525
Middle-aged (RR)	0.120	0.289	0.732
Young adults (GMO)	0.075	0.724	0.362

This table reports the approximation procedure of β's from corresponding hyperbolic discount functions for different age groups estimated in Green et al. *(1999) (GMO) and Read and Read (2004) (RR).*

C Partial Sophistication

In our baseline model, we assume that people are fully sophisticated, meaning all agents in the economy forecast the self-control problems faced by future selves and descendants perfectly. In this section, we analyse whether our results depend on this assumption. We do so by allowing people to be partially sophisticated in the following way. At each age $i \in \{1, ..., I\}$, with probability $(1 - \pi_i) \in [0, 1]$, agent i believes that starting with next period onwards all the future selves and descendants have perfect self-control, and hence, they all discount according to δ discount factor only. With the remaining probability, π_i, agent i knows the true economic environment. Thus, π_i represents the awareness (sophistication) of self i regarding the self-control problems. The vector, $\pi = (\pi_1, \pi_2, .., \pi_I)$, then represents the sophistication profile of an individual over the life cycle. The way we model partial sophistication does not follow the seminal paper of O'Donoghue and Rabin (1999), and is more in line with Eliaz and Spiegler (2006) and Asheim (2007).[1]

We set the partial sophistication model up for the general setup where the economy starts from any initial level of capital stock and prices change over time. We first define the parent's problem under partial sophistication in the general setup.

[1]We justify our way of modelling partial sophistication on the grounds of tractability. The added bonus of our model of partial sophistication is that the structure is consistent with a learning approach to sophistication (e.g., Ali (2011)).

Parent's Problem under Partial Sophistication (along the Transition)

$$
\begin{aligned}
V(b,k;\tau) &= \max_{b_0} \; u\left(R(k)\left(1-\tau_I\right)b + w(k) + T_I - b_0\right) + \\
&\quad +\delta\left[\sum_{i=0}^{I-1} \delta^i u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(b_0)...) + w(k^{i+1}) + T_i - b_{i+1}\right) + \delta^I V\left(b_I(...(b_0)...), k^{I+1};\tau\right)\right]
\end{aligned}
$$

s.t. for all b_0

$$
\begin{aligned}
b_1(b_0,k^1;\tau) = \arg\max_{\hat{b}_1} \; & u\left(R(k^1)\left(1-\tau_0\right)b_0 + w(k^1) + T_0 - \hat{b}_1\right) + \\
+\delta\beta_1 & \left[\begin{array}{c} \pi_1\left\{\sum_{i=1}^{I-1} \delta^{i-1} u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(\hat{b}_1)...) + w(k^{i+1}) + T_i - b_{i+1}(...(\hat{b}_1)...)\right) + \delta^{I-1} V\left(b_I(...(\hat{b}_1)...), k^{I+1};\tau\right)\right\} \\ +\left(1-\pi_1\right)W_1\left(\hat{b}_1, k^2;\tau\right) \end{array}\right]
\end{aligned}
$$

s.t. for all b_1

$$
\begin{aligned}
b_2(b_1,k^2;\tau) = \arg\max_{\hat{b}_2} \; & u\left(R(k^2)\left(1-\tau_1\right)b_1 + w(k^2) + T_1 - \hat{b}_2\right) \\
+\delta\beta_2 & \left[\begin{array}{c} \pi_2\left\{\sum_{i=2}^{I-1} \delta^{i-2} u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(\hat{b}_2)...) + w(k^{i+1}) + T_i - b_{i+1}(...(\hat{b}_2)...)\right) + \delta^{I-2} V\left(b_I(...(\hat{b}_2)...), k^{I+1};\tau\right)\right\} \\ +\left(1-\pi_2\right)W_2\left(\hat{b}_2, k^3;\tau\right) \end{array}\right]
\end{aligned}
$$

s.t. for all b_2

...

s.t. for all b_{I-2}

$$
\begin{aligned}
b_{I-1}(b_{I-2},k^{I-1};\tau) \in \arg\max_{\hat{b}_{I-1}} \; & u\left(R(k^{I-1})\left(1-\tau_{I-2}\right)b_{I-2} + w(k^{I-1}) + T_{I-2} - \hat{b}_{I-1}\right) + \delta\beta_{I-1}\left(1-\pi_{I-1}\right)W_{I-1}\left(\hat{b}_{I-1}, k^I;\tau\right) \\
& +\delta\beta_{I-1}\left[\pi_{I-1}\left\{u\left(R(k^I)\left(1-\tau_{I-1}\right)\hat{b}_{I-1} + w(k^I) + T_{I-1} - b_I(...(\hat{b}_{I-1})...)\right) + \delta V\left(b_I(...(\hat{b}_{I-1})...), k^{I+1};\tau\right)\right\}\right]
\end{aligned} \tag{12}
$$

s.t. for all b_{I-1}

$$
b_I(b_{I-1},k^I;\tau) = \arg\max_{\hat{b}_I} \; u\left(R\left(1-\tau_{I-1}\right)b_{I-1} + w_{I-1} + T_{I-1} - \hat{b}_I\right) + \delta\beta_I\left[\pi_I V\left(\hat{b}_I, k^{I+1};\tau\right) + \left(1-\pi_I\right)W_I\left(\hat{b}_I, k^{I+1};\tau\right)\right] \tag{13}
$$

where the functions W_i for $i = 0, 1, .., I-1$ solve:

$$
W_i\left(b,k;\tau\right) = \max_{b'} \; u\left(R\left(1-\tau_i\right)b + w_i + T_i - b'\right) + \delta W_{i+1}\left(b',k';\tau\right);
$$

with

$$
W_I\left(b,k;\tau\right) = \max_{b'} \; u\left(R\left(1-\tau_I\right)b + w_I + T_I - b'\right) + \delta W_0\left(b',k';\tau\right).
$$

To understand the nested nature of policies and the way we model partial sophistication better, let us analyse the definition of policies in (12) and (13). First, constraint (13) describes how self I chooses b_I. The number $\pi_I \in [0,1]$ represents the belief of self I about the presence of self-control problems. More precisely, this is the belief of self I about the probability that next period when he becomes a parent he will face an offspring with self-control problems, i.e. $(\beta_1, \ldots, \beta_I) \neq (1, \ldots, 1)$, and the offspring will face an offspring with self-control problems, and so on. Note that in reality this probability is one, meaning in each generation people face self-control problems over their life cycle. If $\pi_I < 1$, self I is partially naive in the sense that he incorrectly attaches positive probability $(1-\pi_I)$ to the event that there will never be self-control problems in the future, i.e. $(\beta_1, \ldots, \beta_I) = (1, \ldots, 1)$. So, in our environment, π_I represents the level of sophistication of self I. We assume that all agents, including the parents, correctly guess the level of sophistication of their future selves, $(\pi_i)_i$. In other terms, agents share the same higher-order beliefs.[2] Second, consider constraint (12) which defines how self $I-1$ chooses b_{I-1}. The number $\pi_{I-1} \in [0,1]$ represents the degree of sophistication of self $I-1$, meaning self $I-1$ knows the truth that his followers will have self-control problems with probability π_{I-1}. In particular, with π_{I-1} probability self $I-1$ thinks self I chooses b_I according to (13), and with the remaining probability he thinks self I chooses b_I without facing any self-control problems. We have just seen that the last constraint, (13), enters the parent's problem in at least two ways: first, in the definition of self I's policy function and then as a constraint in the definition of self $I-1$'s policy function. These two different constraints are represented by a single constraint, (13), because the parent and self $I-1's$ sophisticated belief agree about how self I will behave.[3] Similarly, the constraint describing self $I-1$'s policy is also a constraint in the constraint that describes self $I-2$'s policy, and self $I-2$'s policy is also a constraint of self $I-3$'s, and so on. Thus, actually the constraint that describes the policy of self i enters parent's problem in i different places but since these are all identical constraints, we represent them with just one constraint that describes self $i's$ policy.

A Stationary Markov equilibrium *with taxes τ consists of a level of capital k, prices R, w, value functions $V(\cdot;\tau)$ and $\{W_i(\cdot;\tau)\}_{i=0}^{I}$ and policy functions $\{b_i(\cdot;\tau)\}_i$ such that: (i) the prices satisfy (2) in the main text; (ii) the value functions and the policies are consistent with the parent's problem described above; (iii) the government budget is satisfied period-by-period and markets clear: $T_i = R\tau_i b_i(k;\tau)$ and $b_i(k;\tau) = k$ for all i.*

Proposition C.1. below proves that if the constant relative risk aversion coefficient σ is equal to 1, meaning utility is logarithmic, then the degree of sophistication is immaterial for taxes.

Proposition C.1. *Suppose $u(c) = log(c)$. Then, for any level of partial sophistication over the life cycle, π, optimal taxes take the exact form of those in Proposition 1.*

Proof. Relegated to Appendix C.1. ■

The invariance of optimal taxes to the level of sophistication for logarithmic utility is analogous to the equivalence result obtained by Pollak (1968) on consumption policies in a partial equilibrium environment. Proposition C.1. generalises this result to a general equilibrium environment where partial sophistication is modelled differently from O'Donoghue and Rabin (1999) which is the standard model of partial sophistication in the literature.

It is evident from Proposition C.1. that in order to investigate the robustness of our policy findings with respect to naiveté, we need to move away from the assumption of $\sigma = 1$. Unfortunately, when $\sigma \neq 1$ and agents are allowed to be partially sophisticated, we do not get closed form solutions for optimal

[2]Of course, this structure is rich enough to allow for disagreements on higher order beliefs across agents as in O'Donoghue and Rabin (2001). At the same time, if certain regularity conditions are satisfied, it is possible to map such disagreements within a learning environment à la Ali (2011) as either coming from different priors about each other's sophistication or from different information sets across agents. Details are available upon request.

[3]Sophisticated belief of self i about how self j, $j > i$, agrees with parent's belief thanks to our assumption that the same 'beliefs' $(\pi_i)_i$ are shared by all agents.

taxes. Therefore, we have to resort to numerical analysis. For simplicity, we keep the assumption that the economy is at a steady state. The details of our computational procedure are explained in Appendix C.2.

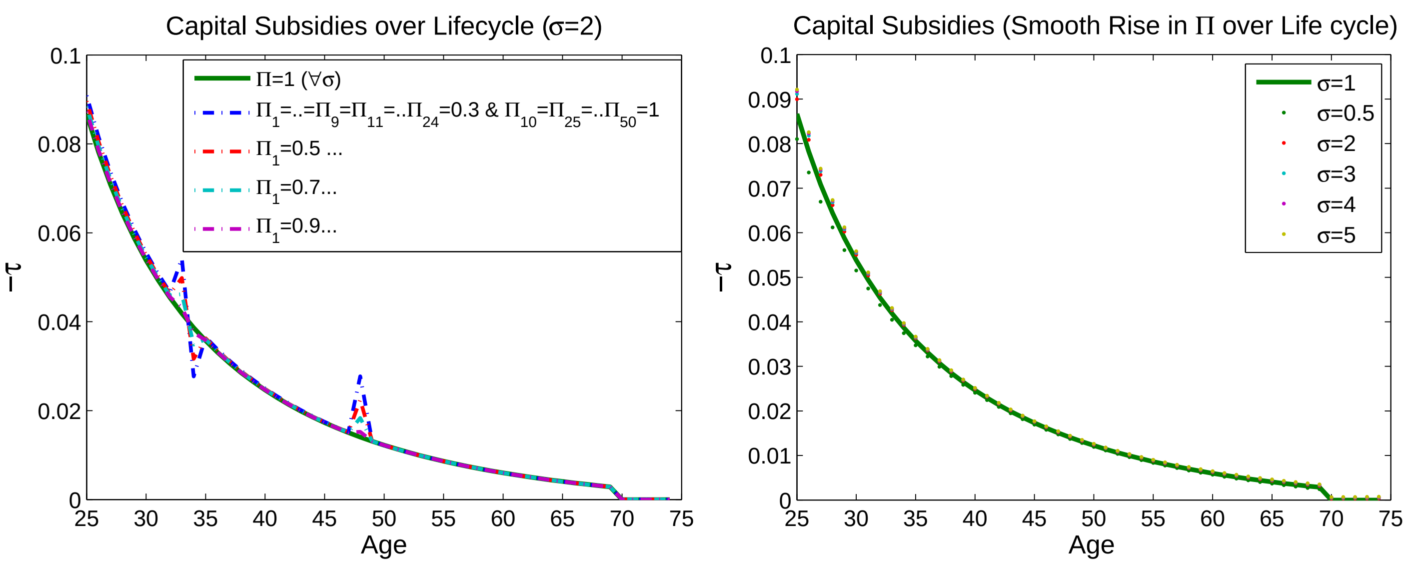
First, we set $\sigma = 2$ and analyse how different patterns of partial sophistication over the life cycle affect optimal subsidies. Throughout this section, we set the life-cycle self-control pattern according to our benchmark calibration, i.e. the first line of Table 1 in the main text. In Figure 1a, the blue solid curve represents the benchmark case of full sophistication, $\pi_i = 1$, for all i. Each dashed curve represents a life-cycle pattern where sophistication level starts at π at the beginning of life and is constant until period 10 when it jumps to 1 and in period 11 it jumps back to π. Then, there is a second jump in period 25, but this is a permanent one: agent remains fully sophisticated from then on. We simulate optimal subsidies for $\pi = 0.3, 0.5, 0.7$, and 0.9, and plot them in Figure 1a with dashed lines. The figure shows that the level of optimal subsidies differ significantly from the benchmark case with full sophistication only in periods which are followed by a sharp change in the level of sophistication in the subsequent period.

The dotted lines in Figure 1b plot optimal subsidies when the level of sophistication changes smoothly over the life cycle for various values of σ.[4] The solid blue line again represents the fully sophistication benchmark (under any σ because the steady-state condition holds). This figure first of all confirms the previous finding: the degree of sophistication does not matter for optimal subsidies as long as there are no abrupt changes in sophistication. Figure 1b also suggests that, as σ moves away from 1, the effect of sophistication becomes more significant. However, even when $\sigma = 5$, the difference between optimal subsidies under full sophistication (the blue line) and the partially sophisticated model is around 0.05% for the first period and this difference decreases to below 0.01% after the fourth period. Finally, in Figure 1b, the optimal subsidies under partial sophistication for $\sigma = 0.5$ are depicted by the dotted line that lies below the full sophistication line whereas the subsidies for all $\sigma > 1$ are depicted by the dotted lines that lie above it. This observation suggest a qualitative pattern: that for $\sigma > 1 (< 1)$, optimal taxes increase (decrease) with the level of sophistication.[5]

We conclude that, as long as the level of naiveté is not changing abruptly from one period to another, the level optimal capital subsidies over the life cycle is robust to various scenarios about how sophistication changes with age. Moreover, when the level of partial sophistication is changing smoothly (or constant), the level optimal capital subsidies over the life cycle is not significantly affected by our choice of the coefficient of constant relative risk aversion.

[4] To be precise, sophistication depends on age according to the concave function $\pi(i) = [1 - \frac{3(1+I-i)}{4I}]^{1/2}$.

[5] An earlier related result is given in O'Donoghue and Rabin (2003) which shows that, when we model partial sophistication a la O'Donoghue and Rabin (1999), if $\sigma > 1(< 1)$, then more sophisticated people over-consume less (more). O'Donoghue and Rabin (2003) does not analyze taxes but the tax implication of their finding is obvious: if $\sigma > 1(< 1)$, then more sophisticated people should be taxed more (less) heavily. We have shown that this result is valid under our way of modelling partial sophistication as well. The derivations are available upon request.



(a) *Jumps in π* (b) *Smoothly rising π*

Figure 1: *Partial Sophistication*

C.1 Proof of Proposition C.1.

The proof follows closely the proof of Proposition 1. Letting b_i and k^{i+1} be the saving level in period i and aggregate capital stock in period $i+1$, define $\Gamma_i(b_i, k^{i+1})$ and $G_i(k^{i+1})$ as in the proof of Proposition 1. Similarly, define $c_{i+1}(b_i, k^{i+1}) = M_{i+1}\Gamma_i(b_i, k^{i+1})$.

Now using linearity of the policy functions and the first-order approach, we can rewrite the parent's problem as:

$$V(b,k;\tau) = \max_{M_0} \; u\left(M_0\Gamma_I(b)\right) + \delta\left[\sum_{i=0}^{I-1} \delta^i u\left(Q_i(k)M_{i+1}\Gamma_I(b)\right) + \delta^I V\left((1-M_I)Q_{I-1}(k)\Gamma_I(b), k^{I+1};\tau\right)\right] \quad (14)$$

$$s.t. \quad \text{for all } i \in \{1, ..., I-1\}$$

$$(M_iQ_{i-1}(k)\Gamma_I(b,k))^{-\sigma} = \delta\beta_i \left[\begin{array}{c} \pi_i R(k^{i+1})(1-\tau_i(k^{i+1}))\left\{\begin{array}{c} \sum_{j=i+1}^{I} \delta^{j-(i+1)}\left(M_jQ_{j-1}(k)\Gamma_I(b,k)\right)^{-\sigma} M_j \frac{Q_{j-1}(k)}{Q_i(k)} \\ +\delta^{I-i}V'(b_I(..(b)..), k^{I+1};\tau)(1-M_I)\frac{Q_{I-1}(k)}{Q_i(k)} \end{array}\right\} \\ +(1-\pi_i)W_i'\left(b_i(..(b)..), k^{i+1};\tau\right) \end{array}\right]$$

$$(M_IQ_{I-1}(k)\Gamma_I(b,k))^{-\sigma} = \delta\beta_I\left[\pi_I V'\left(b_I(..(b)..), k^{I+1};\tau\right) + (1-\pi_I)W_I'\left(b_I(..(b)..), k^{I+1};\tau\right)\right].$$

Core proof of Proposition C.1.

We will prove that facing the sequence of efficient capital levels and the taxes specified in Proposition 1, people will choose the efficient allocation, thereby verifying both (1) that the sequence of the efficient capital levels is actually part of equilibrium under the taxes described in Proposition 1, and (2) that under the taxes specified by Proposition 1, people choose the efficient allocation.

Guess

$$\begin{aligned} V(b,k;\tau) &= D\log(\Gamma_I(b,k)) + B(k), \\ W_i(b,k;\tau) &= D_i\log(\Gamma_i(b,k)) + B_i(k), \text{ for } i = 0,..,I \end{aligned}$$

where D and $D_0, D_1, .., D_I, B_0, ..., B_I$ are constants of the parent's and naive self-i's value functions.

STEP 1: Compute the coefficients for the naive value functions, $D_0, .., D_I$.

If we let $k' = K(k)$, from the first-order condition for the W_i problem, we have (after tedious calculations):

$$b_i(b,k;\tau) = \frac{R(k)(1-\tau_i(k))b + w(k) + T_i(k) - [G_{i+1}(k') + w(k') + T_{i+1}(k')]\left[\delta R(k')(1-\tau_{i+1}(k'))D_{i+1}\right]^{-1}}{1 + \left[\delta R(k')(1-\tau_{i+1}(k'))D_{i+1}\right]^{-1} R(k')(1-\tau_{i+1}(k'))}.$$

Plugging this in the value function, and performing some tedious re-arrangements, we get for $i = 0, 1, .., I$:

$$D_i = (1 + \delta D_{i+1})$$

and

$$D_I = (1 + \delta D_0).$$

Thus,

$$D_0 = D_1 = .. = D_I = \frac{1}{1-\delta}.$$

STEP 2: Compute the coefficients for parent's value function, D.

Take $D_1, .., D_I$ from above. Compute V' and W_i' for $i = 0, 1, .., I$ in terms of D, D_i using the guesses for value functions:

$$\begin{aligned} V'(b_I(..(b)..), k^{I+1};\tau) &= DR(k^{I+1})(1-\tau_I(k^{I+1}))(\Gamma_I(b,k)Q_I(k))^{-1}, \\ W_i'(b_i(..(b)..), k^{i+1};\tau) &= D_iR(k^{i+1})(1-\tau_i(k^{i+1}))(\Gamma_I(b,k)Q_i(k))^{-1}, \end{aligned} \tag{15}$$

where we used the recursion (1).

Plugging these in the constraints described in problem (14), we get for all $i \in \{1, ..., I-1\}$:

$$\begin{aligned} (M_iQ_{i-1}(k))^{-1} &= \delta\beta_iR(k^{i+1})(1-\tau_i(k^{i+1}))\,(Q_i(k))^{-1}\left[\begin{array}{c} \pi_i\left\{\sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\right\} \\ +(1-\pi_i)\,D_i \end{array}\right] \\ & \qquad \textit{and} \\ (M_IQ_{I-1}(k))^{-1} &= \delta\beta_IR(1-\tau_I(k^{I+1}))\,(Q_I(k))^{-1}\left[\pi_ID + (1-\pi_I)\,D_I\right]. \end{aligned}$$

Now, using the marginal condition describing self-I behaviour, it is easy to show that

$$M_I(D) = \frac{1}{1 + \beta_I\delta(\pi_ID + (1-\pi_I)\,D_I)}.$$

Similarly, use other constraints defining the policies to compute $M_i(D)$ for $i = 1, .., I-1$:

$$M_i(D) = \frac{1}{1 + \beta_i\delta\left(\pi_i\left\{\sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\right\} + (1-\pi_i)\,D_i\right)}.$$

Taking first-order condition with respect to bequests in the parent's problem (14) and plugging in the $M_i(D)$ from above, we get:

$$M_0(D) = \frac{1}{1 + \delta\left(\sum_{j=0}^{I-1}\delta^j + \delta^I D\right)}.$$

Now verify the value function to compute D :

$$\begin{aligned} D\log\left(\Gamma_I\left(b,k\right)\right) + B(k) &= \log\left(M_0(D)\Gamma_I\left(b,k\right)\right) \\ &+\delta\left[\sum_{i=0}^{I-1}\delta^i\log\left(Q_i(k)M_{i+1}(D)\Gamma_I\left(b,k\right)\right) + \delta^I\left\{D\log\left(\Gamma_I\left(b,k\right)Q_I(k)\right) + B(k^{I+1})\right\}\right], \end{aligned}$$

which implies

$$D = \sum_{i=0}^{I}\delta^i + \delta^{I+1}D$$

and hence

$$D = \frac{1}{1-\delta}.$$

By plugging D in the formula for $M_i(D)$, we compute

$$\begin{aligned} M_i &= \frac{1-\delta}{1-\delta+\beta_i\delta},\ \text{for all } i \in \{1, .., I\}, \\ M_0 &= 1-\delta. \end{aligned} \tag{16}$$

Now we turn to taxes that implement the efficient allocation. The constraint that describes self-i's behaviour for $i \in \{1, .., I-1\}$ becomes the following once we plug in the derivatives of the value functions from (15) :

$$\left(M_iQ_{i-1}(k)\Gamma_I\left(b,k\right)\right)^{-1} = \delta\beta_iR(k^{i+1})(1-\tau_i(k^{i+1}))\left(M_{i+1}Q_i(k)\Gamma_I\left(b,k\right)\right)^{-1}\left[\begin{array}{c}\pi_i\left\{\sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\right\} \\ +(1-\pi_i)\,D_i\end{array}\right]M_{i+1}. \tag{17}$$

The comparison of (17) with the efficiency condition (1) in the main paper gives the optimal tax as:

$$\begin{aligned} \left(1-\tau_i^*(k^{i+1})\right) &= \frac{1}{\beta_i}\left(\left[\begin{array}{c}\pi_i\left\{\sum_{j=i+1}^{I}\delta^{j-(i+1)} + \delta^{I-i}D\right\} \\ +(1-\pi_i)\,D_i\end{array}\right]M_{i+1}\right)^{-1} \\ &= \frac{1}{\beta_i}\left(1-\delta+\beta_{i+1}\delta\right), \end{aligned}$$

where from the first to the second equality we used (16). For self-I, the constraint describing his behaviour in problem (14) reads as follows:

$$(M_I Q_{I-1}(k)\Gamma_I(b,k))^{-1} = \delta\beta_I R(k^{I+1})(1-\tau_I(k^{I+1}))\,(M_0 Q_I(k)\Gamma_I(b,k))^{-1}[\pi_I D + (1-\pi_I)D_I]\,M_0,$$

and the comparison of this with the efficiency condition gives

$$\left(1-\tau_I^*(k^{I+1})\right) = \frac{1}{\beta_I}.$$

Finally, a comparison of the following first-order condition of the parent

$$(M_0\Gamma_I(b,k))^{-1} = \delta R(k^1)(1-\tau_0(k^1))(M_1 Q_0(k)\Gamma_I(b,k))^{-1}\frac{\left[\sum_{i=0}^{I-1}\delta^i + \delta^I D\right]}{M_1^{-1}}$$

with the corresponding optimality condition gives

$$1-\tau_0^*(k^1) = (1-\delta+\beta_1\delta).$$

C.2 Computational Procedure

C.2.1 Guess:

Guess[6]

$$\begin{aligned} V(b;\tau) &= D(\tau)\frac{(\Gamma_I(b))^{1-\sigma}}{1-\sigma}, \\ W_i(b;\tau) &= D_i(\tau)\frac{(\Gamma_i(b))^{1-\sigma}}{1-\sigma}, \end{aligned}$$

where D and D_i for $i = 0,1,..,I$ are constants of the parent's and naive self-i's value functions. Observe that these constants depend on the tax system, τ. In what follows, for notational simplicity this dependence will be implicit.

C.2.2 Characterising equilibrium value function constants for a given tax system τ:

STEP 1: Computing equilibrium $D_0,..,D_I$.

From the first-order conditions for the W_i problem, we have: for all $i \in \{0,1,..,I-1\}$

$$\begin{aligned} D_i &= \left[\frac{[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1}{\sigma}}R(1-\tau_{i+1})}{1+[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1}{\sigma}}R(1-\tau_{i+1})}\right]^{1-\sigma}\left(1+\delta\frac{D_{i+1}}{[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1-\sigma}{\sigma}}}\right), \qquad (18) \\ D_I &= \left[\frac{[\delta R(1-\tau_0)D_0]^{-\frac{1}{\sigma}}R(1-\tau_0)}{1+[\delta R(1-\tau_0)D_0]^{-\frac{1}{\sigma}}R(1-\tau_0)}\right]^{1-\sigma}\left(1+\delta\frac{D_0}{[\delta R(1-\tau_0)D_0]^{-\frac{1-\sigma}{\sigma}}}\right). \end{aligned}$$

Given taxes, the solution to these $I+1$ equations give us $I+1$ unknowns, $D_0,..,D_I$.

[6] The `matlab` files used for all our simulations are available online. We refer to the README.pdf file for details.

STEP 2: Computing equilibrium D.

From our guess of the value function, we have

$$V'(b_I;\tau) = D(\Gamma_I(b_I))^{-\sigma} R(1-\tau_I),$$

and by envelope we have

$$V'(b_I;\tau) = R(1-\tau_I)u'(c_0) = R(1-\tau_I)\left(M_0\Gamma_I(b_I)\right)^{-\sigma},$$

which together imply

$$D = M_0^{-\sigma}. \tag{19}$$

C.2.3 Characterising optimal tax system, τ^*:

The incentive constraints for agents $i = 1, ..I$ together with parent's optimality condition with respect to bequest decision characterise the solution to the parent's problem and hence the equilibrium for a given tax system, τ. Comparison of these $I+1$ equations with the corresponding commitment Euler equations, we immediately see that optimal taxes must satisfy:

$$\begin{aligned}
\text{For all } i &\in \{0, .., I-2\}, \\
\left(1-\tau^*_{i+1}\right) &= \frac{1}{\beta_{i+1}}\left(\frac{\left[\pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{i-(i+2)}\left(M_j^*\frac{Q^*_{j-1}}{Q^*_{i+1}}\right)^{1-\sigma} + \delta^{I-(i+1)}D^*\left(\frac{Q^*_I}{Q^*_{i+1}}\right)^{1-\sigma} + \left(1-\pi_{i+1}\right)D^*_{i+1}\right\}\right]}{M_{i+2}^{*-\sigma}}\right)^{-1} \\
\left(1-\tau^*_I\right) &= \frac{1}{\beta_I}\left(\frac{\left[\pi_I D^* + \left(1-\pi_I\right)D^*_I\right]}{M_0^{*-\sigma}}\right)^{-1} \\
\left(1-\tau^*_0\right) &= \left(\frac{\left[\sum_{i=1}^{I}\delta^{i-1}\left(M_i\frac{Q^*_{i-1}}{Q^*_0}\right)^{1-\sigma} + \delta^I D^*\left(\frac{Q^*_I}{Q^*_0}\right)^{1-\sigma}\right]}{M_1^{*-\sigma}}\right)^{-1},
\end{aligned} \tag{20}$$

where D^* and D_i^* are the values associated with the efficient allocation computed according to (19) and (18) evaluated at the optimal taxes.

C.2.4 Iteration

1. Before starting the iteration, compute efficient consumption and saving allocations $(c_i^*, b_i^*)_{i=0}^{I}$ according to:

$$\begin{aligned}
c_0^* &= Rb\frac{(R^{I+1}-1)}{R^{I+1}}\frac{1}{\sum_{i=0}^{I}\left(\frac{(R\delta)^{\frac{1}{\sigma}}}{R}\right)^i}, \\
\text{for all } i &\in \{0, .., I-1\},\ c^*_{i+1} = c_i^*(R\delta)^{\frac{1}{\sigma}}, \\
b_0^* &= Rb - c_0^*, \\
\text{for all } i &\in \{0, .., I-1\},\ b^*_{i+1} = Rb_i^* - c^*_{i+1}.
\end{aligned}$$

2. Start with a guess for the efficient tax system $\tau = (\tau_0, .., \tau_I)$, where is given by government's period budget constraint $T_i = Rb_i^* \tau_i$ (for the initial guess we use optimal taxes in the logarithmic case).

3. Compute the linear policy functions according to formulas:

$$\begin{aligned} M_0 &= \frac{c_0^*}{Rb(1-\tau_I) + T_I + G_I} = \frac{c_0^*}{Rb + G_I}, \\ \text{For all } i &\in \{0, 1, .., I-1\},\ M_{i+1} = \frac{c_{i+1}^*}{Rb_i^*(1-\tau_i) + T_i + G_i} = \frac{c_{i+1}^*}{Rb_i^* + G_i}, \end{aligned}$$

where

$$G_I = \frac{1}{1 - \left[R^{I+1} \prod_{j=0}^{I} (1-\tau_j)\right]^{-1}} \sum_{i=0}^{I} \frac{T_i + w}{R^{i+1} \prod_{j=0}^{i} (1-\tau_j)}$$

and for all $i \in \{0, .., I-1\}$

$$G_i = \frac{G_{i+1} + Rb_{i+1}^* \tau_{i+1} + w}{R(1-\tau_{i+1})}.$$

4. Compute D and $D_1, ..D_I$ according to (19) and (18) evaluated at the tax guess.

5. Now use the linear policies computed in step 3 and the value function constants computed in step 4 to compute taxes according to the system of equations describing optimal taxes (20).

6. If the taxes you compute in step 5 is the same as the taxes you started the last iteration, stop. If not, use the taxes you computed in step 5 as the new guess and continue iteration.

D Introducing an Illiquid Asset

To simplify our analysis, consider a three period version of our model. With one difference: there is an additional asset people can buy in period one. Also, again for simplicity, we assume $\beta_1 = 0$. This asset, denoted by d_1, is illiquid in the sense that it does not pay in period two, but pays in period 3 an after tax return $R^d(1-\tau^d)d_1$. Self 2's problem then is:

$$\begin{aligned} c_2, c_3 &\in \arg\max_{c_2, c_3} u(c_2) + \bar{\beta}_2 \delta u(c_3) \\ & \quad s.t. \\ c_2 + \frac{c_3}{R(1-\tau_2)} &\leq R(1-\tau_1)b_1 + T_1 + \frac{T_2}{R(1-\tau_2)} + \frac{R^d(1-\tau^d)d_1}{R(1-\tau_2)} \equiv y_1(b_1, d_1) \end{aligned}$$

Let $c_2(y_1), c_3(y_1)$ be the solution to the above problem when $\bar{\beta}_2 = \beta_2$ and $\hat{c}_2(y_1), \hat{c}_3(y_1)$ when $\bar{\beta}_2 = 1$.

Self 1's problem:

$$\begin{aligned} &\max_{b_1, d_1} u(k_0 - b_1 - d_1) + \pi_1 \delta \left[u(c_2(y_1)) + \delta u(c_3(y_1))\right] \\ &+ (1-\pi_1)\delta \left[u(\hat{c}_2(y_1)) + \delta u(\hat{c}_3(y_1))\right]. \end{aligned}$$

Case 1. Government sets taxes such that

$$R^d(1-\tau^d) < R^2(1-\tau_1)(1-\tau_2).$$

In this case, obviously $d_1 = 0$. So, it is as if there are no illiquid assets; government prevents people from using these assets through taxes. Then, simply by setting τ_1, τ_2 exactly equal to the efficient taxes in the environment without illiquid asset, τ_1^*, τ_2^*, we implement the efficient allocation in the market with the illiquid asset. Let us compute these taxes for future use. Since

$$u'(c_2) = \beta_2 \delta R(1-\tau_2)u'(c_3),$$

efficiency requires

$$(1-\tau_2^*) = \frac{1}{\beta_2}.$$

To compute optimal period one tax, take first-order condition of the parent's problem with respect to b_1 :

$$u'(c_1) = \delta \left(\begin{array}{c} \pi_1 \left[u'(c_2(y_1))c_2'(y_1)\frac{\partial y_1(b_1,d_1)}{\partial b_1} + \delta u'(c_3(y_1))c_3'(y_1)\frac{\partial y_1(b_1,d_1)}{\partial b_1}\right] \\ +(1-\pi_1)\left[u'(\hat{c}_2(y_1))\hat{c}_2'(y_1)\frac{\partial y_1(b_1,d_1)}{\partial b_1} + \delta u'(\hat{c}_3(y_1))\hat{c}_3'(y_1)\frac{\partial y_1(b_1,d_1)}{\partial b_1}\right] \end{array} \right)$$

where $\frac{\partial y_1(b_1,d_1)}{\partial b_1} = R(1-\tau_1)$ (For ease of exposition, assume the policies are differentiable).[7] Therefore,

$$u'(c_1) = \delta R(1-\tau_1) \left(\begin{array}{c} \pi_1 \left[u'(c_2(y_1))c_2'(y_1) + \delta u'(c_3(y_1))c_3'(y_1)\right] \\ +(1-\pi_1)\left[u'(\hat{c}_2(y_1))\hat{c}_2'(y_1) + \delta u'(\hat{c}_3(y_1))\hat{c}_3'(y_1)\right] \end{array} \right)$$

which implies:

$$(1-\tau_1^*) = \frac{u'(c_1^*)}{\delta R\left(\pi_1\left[u'(c_2^*)c_2'(y_1^*) + \delta u'(c_3^*)c_3'(y_1^*)\right] + (1-\pi_1)\left[u'(\hat{c}_2^*)\hat{c}_2'(y_1^*) + \delta u'(\hat{c}_3^*)\hat{c}_3'(y_1^*)\right]\right)},$$

where y_1^* is the net present value of wealth under the efficient allocation.

Case 2. Government sets taxes such that

$$R^d(1-\tau^d) \geq R^2(1-\tau_1)(1-\tau_2).$$

Then, obviously, agents might be using $d_1 \geq 0$. In that case, since

$$u'(c_2) = \beta_2 \delta R(1-\tau_2)u'(c_3)$$

still holds, efficiency still requires

$$(1-\tau_2^*) = \frac{1}{\beta_2}.$$

[7] It is well-known that in general we cannot guarantee even the continuity of the policy functions (e.g., see Krusell and Smith (2003), and Harris and Laibson (2001)).

To see optimal taxes on the illiquid asset, consider the first-order condition with respect to d_1 :

$$u'(c_1) = \delta \left(\begin{array}{c} \pi_1 \left[u'(c_2(y_1)) c_2'(y_1) \frac{\partial y_1(b_1,d_1)}{\partial d_1} + \delta u'(c_3(y_1)) c_3'(y_1) \frac{\partial y_1(b_1,d_1)}{\partial d_1} \right] \\ +(1-\pi_1) \left[u'(\hat{c}_2(y_1)) \hat{c}_2'(y_1) \frac{\partial y_1(b_1,d_1)}{\partial d_1} + \delta u'(\hat{c}_3(y_1)) \hat{c}_3'(y_1) \frac{\partial y_1(b_1,d_1)}{\partial d_1} \right] \end{array} \right)$$

where $\frac{\partial y_1(b_1,d_1)}{\partial d_1} = \frac{R^d(1-\tau^d)}{R(1-\tau_2)}$. Therefore,

$$u'(c_1) = \delta \frac{R^d(1-\tau^d)}{R(1-\tau_2)} \left(\begin{array}{c} \pi_1 \left[u'(c_2(y_1)) c_2'(y_1) + \delta u'(c_3(y_1)) c_3'(y_1) \right] \\ +(1-\pi_1) \left[u'(\hat{c}_2(y_1)) \hat{c}_2'(y_1) + \delta u'(\hat{c}_3(y_1)) \hat{c}_3'(y_1) \right] \end{array} \right)$$

which implies:

$$\begin{aligned} & R^d(1-\tau^{d*}) = \\ & = R(1-\tau_2^*) \frac{u'(c_1^*)}{\delta R \left(\pi_1 \left[u'(c_2^*) c_2'(y_1^*) + \delta u'(c_3^*) c_3'(y_1^*) \right] + (1-\pi_1) \left[u'(\hat{c}_2^*) \hat{c}_2'(y_1^*) + \delta u'(\hat{c}_3^*) \hat{c}_3'(y_1^*) \right] \right)} \\ & = R(1-\tau_2^*) R(1-\tau_1^*). \end{aligned} \tag{21}$$

As a result, when there is an illiquid asset, government can either prevent people from using this asset by taxing it heavily or has to tax it according to (21). In either case, the taxes on period one and period two liquid assets are exactly equal to the optimal taxes in the environment without illiquid assets.

Nicola Pavoni, Department of Economics, Bocconi University, via Roentgen 1, 20136 Milan, Italy

Hakki Yazici, Sabanci University, FASS, Orta Mahalle, Universite Caddesi No:27 Tuzla, 34956 Istanbul, Turkey

References

Ali, S.N. (2011). 'Learning self-control', *The Quarterly Journal of Economics*, vol. 126(2), pp. 857–893.

Asheim, G.B. (2007). 'Procrastination, partial naivete, and behavioral welfare analysis', Oslo University, Department of Economics.

Eliaz, K. and Spiegler, R. (2006). 'Contracting with diversely naive agents', *Review of Economic Studies*, vol. 73(3), pp. 689–714.

Green, L., Myerson, J. and Ostaszewski, P. (1999). 'Discounting of delayed rewards across the life-span: Age differences in individual discounting functions', *Behavioural Processes*, vol. 46(1), pp. 89–96.

Harris, C.J. and Laibson, D. (2001). 'Dynamic choices of hyperbolic consumers', *Econometrica*, vol. 69(4), pp. 935–957.

Krusell, P. and Smith, A.A. (2003). 'Consumption–savings decisions with quasi–geometric discounting', *Econometrica*, vol. 71(1), pp. 365–375.

Laibson, D., Repetto, A. and Tobacman, J. (2007). 'Estimating discount functions with consumption choices over the lifecycle', National Bureau of Economic Research.

O'Donoghue, T. and Rabin, M. (1999). 'Doing it now or later', *American Economic Review*, vol. 89(1), pp. 103–124.

O'Donoghue, T. and Rabin, M. (2001). 'Choice and procrastination', *The Quarterly Journal of Economics*, vol. 116(1), pp. 121–160.

O'Donoghue, T. and Rabin, M. (2003). 'Self awareness and self control', in (*Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*pp. 217–243, Russell Sage Foundation Press.

Pollak, R.A. (1968). 'Consistent planning', *The Review of Economic Studies*, vol. 35(2), pp. 201–208, ISSN 00346527.

Read, D. and Read, N. (2004). 'Time discounting over the lifespan', *Organizational Behavior and Human Decision Processes*, vol. 94(1), pp. 22–32.