INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Optimal Life-cycle Capital Taxation under Self-Control Problems

Nicola Pavoni and Hakki Yazici

Working Paper n. 467

This Version: September 17, 2012

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

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

Nicola Pavoni[†] and Hakki Yazici[‡]

September 17, 2012

Abstract

We study optimal taxation of savings in an economy where agents face self-control problems and are allowed to be partially naive. We assume that the severity of self-control problems changes over the life-cycle. We focus on quasi-hyperbolic discounting with constant elasticity of intertemporal substitution utility functions and linear Markov equilibria. We derive explicit formulas for optimal taxes that implement the efficient allocation. We show that if agents' ability to self-control increases concavely with age, then savings should be subsidized and the subsidy should decrease with age. We also show that allowing for age-dependent self-control problems creates large effects on the level of optimal subsidies, while optimal taxes are not very sensitive to the level of sophistication.

JEL classification: E21, E62, D03.

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

1 Introduction

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 intertemporal preferences: How an individual feels about a given intertemporal tradeoff is independent of when he is asked. However, laboratory and field studies on intertemporal 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 intertemporal 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]

*We would like to thank Per Krusell, 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.

[†]Bocconi University, IGIER, IFS, and CEPR.

[‡]Sabanci University.

[1]See DellaVigna (2009) for a survey of field studies and Frederick, Loewenstein, and O'Donoghue (2002) for a survey of experimental studies. Also, see Laibson, Repetto, and Tobacman (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).

In this paper, we study optimal capital income taxation over the life-cycle in the presence of self-control problems. A common modeling assumption in the literature is that the degree of self-control problem is constant over time. Even though this assumption might be a good approximation of reality for analyzing many questions, a significant body of empirical studies points to the opposite: like many other personality traits, people's ability to self-control changes as they age. The first set of evidence for changing level of self-control over the life-span comes from personality psychology. As Ameriks, Caplin, Leahy, and Tyler (2007) states "personality psychologists associate self-control with conscientiousness, one of the 'big five' personality factors."[3] 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.[4] Indeed, in a survey article on personality development in adulthood, Caspi, Roberts, Robins, and Trzesniewski (2003) conclude "it appears that the increase in conscientiousness is one of the most robust patterns in personality development, especially in young adulthood." There is a second set of more direct evidence in favor of changing self-control: research on intertemporal 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.[5] All this evidence suggests that, in order to study capital taxation over the life-cycle, one should extend the traditional models of self-control to allow for varying degrees of self-control problem over the life-cycle. This is exactly what the current paper does.

In our model, at all ages, agents make consumption and savings decisions facing self-control problems. 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), which builds on the seminal works of Strotz (1955) and Phelps and Pollak (1968). We extend the Laibson (1997) model in two ways that are important to our analysis. First, we allow for the degree of self-control problem to change over time. Second, we introduce partial sophistication which essentially amounts to allowing for different degrees of self awareness about the existence of the self-control problems. In this environment, we define the first-best allocation as the allocation that would arise in the absence of self-control problems. The main exercise in this paper is to examine the optimal tax policy that implements the first-best allocation. It is well-known that in models of quasi-hyperbolic discounting there is multiplicity of equilibria.[6] We restrict attention to linear Markov equilibria, which are not necessarily unique even under CEIS assumption. In our environment however, since when facing linear future policies each agent's problem is strictly concave, the linear equilibrium of finite-period economies - when it exists - is unique.

[3]Actually, Ameriks, Caplin, Leahy, and Tyler (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."

[4]For example, see John, Gosling, Potter, and Srivastava (2003) and Helson, Jones, and Kwan (2002). Ameriks, Caplin, Leahy, and Tyler (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.

[5]Green, Fry, and Myerson (1994), Green, Myerson, and Ostaszewski (1999), Read and Read (2004), and Ameriks, Caplin, Leahy, and Tyler (2007).

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

We restrict attention to logarithmic utility and 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 is a potentially important message since it 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 - based on the assumption of constant self-control over age - that self-control problems always imply savings subsidies.

When utility is logarithmic, optimal taxes are independent of how sophistication changes over the life-cycle. Moreover, if the economy is in the steady-state and agents are fully sophisticated, then optimal taxes are independent of the CEIS coefficient. These results make the tax formulas computed for the logarithmic case quite general. Using these formulas, we prove that if, as strongly suggested by personality psychologists, the degree of self-control increases with age, then capital should indeed be *subsidized* in all periods. We put forth empirical evidence that suggests that the degree of self-control increases concavely with age. We prove that, if this is the case, then optimal capital subsidies should *decrease* with age. We also study the quantitative effects of age-dependent self-control and find that optimal taxes in our environment are much larger than those implied by models with constant self-control, especially for agents in the early stage of the life-cycle.

Finally, we know from O'Donoghue and Rabin (1999) that allowing for even constant level of partial naivete can change people's behavior.[7] We analyze how changing naivete over the life-span alters our optimal taxation results. When CEIS coefficient is different from one and agents are allowed to be partially sophisticated, closed form solutions for optimal taxes are unavailable. Therefore, we resort to numerical analysis at the steady state. We derive two conclusions from our numerical experiments. First, as long as the level of sophistication is not changing abruptly from one period to another, the pattern of optimal capital subsidies over the lifecycle is surprisingly robust to partial sophistication. Second, this result approximately holds for a large range of CEIS coefficients, which implies that the pattern of optimal capital subsidies is somewhat robust to different levels of CEIS coefficients, at least in the steady state.

Related Literature. Krusell, Kuruscu, and Smith (2002) and Krusell, Kuruscu, and Smith (2010) are the two most closely related papers to the current study. The first one, Krusell, Kuruscu, and Smith (2002), considers infinitely living agents facing self-control problems in the form of quasi-hyperbolic discounting à la Laibson. They restrict attention to logarithmic utility and show that a constant subsidy to investment (similar to our capital subsidy) implements the commitment allocation. Krusell, Kuruscu, and Smith (2010) analyzes optimal taxation in an economy where agents live finitely many periods and have temptation and self-control problems à la Gul and Pesendorfer (2001). They first prove that under CEIS preferences, as the parameter that controls temptation goes to infinity, the optimal policy prescriptions of the quasi-hyperbolic model and the temptation model become identical. Then, they show that for the logarithmic

[7] See Ariely and Wertenbroch (2002) for behavioral evidence on partial sophistication.

utility case, this equivalence result holds for any temptation level, and they compute optimal savings taxes. They show that savings should be subsidized and that this subsidy should be increasing with time due to finite life time effect.[8] Our work differs from these papers along two main dimensions. First and foremost, we allow for changing level of self-control problems over the life-cycle while both papers assume the level of self-control problem to be constant over time. By assuming empirically plausible patterns of self-control problems over the life-cycle, we show that capital subsidies should actually be decreasing with age. Second, we allow for agents to be partially aware of their future self-control problems (partial sophistication) as opposed to assuming people at all ages predict their future self-control level perfectly which is the assumption these papers make. This allows us to study the affects of sophistication on capital subsidies.

As discussed above, an immediate implication of our results is that capital income taxes should be age-dependent. The age-dependence result is also a feature of two sets of earlier contributions that analyze 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., Kocherlakota (2010)). The forces generating age-dependence in the current paper are however completely different from the forces in these papers.

2 Model

The economy is populated by a continuum of a unit measure of dynasties who live for a countable infinity of periods, $t = 1, 2, ...,$ 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 the offspring, knowing that they are going to be replaced by their offsprings next period. People are altruistic and they anticipate their offspring's self-control problems.[9] We use quasi-hyperbolic discounting formalized by Laibson (1997) to model self-control problems as follows.

An agent who is in his ultimate period of life (we refer to this agent as parent from now on) has the following preferences over dynastic consumption stream:

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

where c_0 is the consumption level of the current parent, c_i is the consumption level of the offspring at age i, and c_0' is the

[8] In the infinite horizon version of their model, the subsidies would be constant.

[9] In this paper, we are only interested in analyzing life-cycle capital taxation under self-control problems. Therefore, we could have even assumed there are no intergenerational links and hence no bequest motive. We do model altruism (and assume altruism factor is equal to the time discount factor) to abstract away from the effects of finite life time on taxes (see Krusell, Kuruscu, and Smith (2010) for finite life time effects). In a separate paper, Pavoni and Yazici (2012), we analyze bequest taxation under self-control problems in a much more general environment.

consumption level of the offspring when he becomes a parent. u is the instantaneous utility function and δ refers to both the discount factor and the altruism factor. The offspring has different preferences at different periods of his life:

$$u(c_1) + \beta_1\delta\left[u(c_2) + \delta u(c_3) + ... + \delta^{I-1}u(c_0') + ...\right],$$
$$u(c_2) + \beta_2\delta\left[u(c_3) + ... + \delta^{I-2}u(c_0') + ...\right],$$
$$.$$
$$.$$
$$u(c_I) + \beta_I\delta\left[u(c_0') + ...\right].$$

The first equation above is the agent's preference during his first period of adult life, second equation is his preference during his second period, and so on. 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. If we were to take $\beta_i = \beta$ for all i, as previous papers have assumed, that would mean that the degree of self-control problem is constant as people age. However, as documented by personality psychologists and experimental studies, as people age, the severity of the self-control problem they face might change. Therefore, we allow for the severity of self-control problems, β_i, to depend on i.

Another dimension of self-control problems is the extent to which agents can predict the level of self-control problems their followers (be it their future selves or their offsprings) face. We allow for partial sophistication which essentially amounts to allowing for different degrees of self awareness about the existence of self-control problems.[10] We explain in detail the way we model partial sophistication in the next subsection.

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

$$f(k) = F(k,1) + (1-d)k,$$

[10] We are not the first ones to model partial sophistication, O'Donoghue and Rabin (1999) is. However, the way we introduce partial sophistication is different from theirs. We justify our way of modeling partial sophistication on the grounds of tractability and the fact that the two models deliver very similar predictions. The other added bonus of our model partial sophistication is that the structure is consistent with a learning approach to sophistication (e.g., Ali (2011)).

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 all agents are identical, so are their asset holdings. Let b_t be the amount of asset holdings of the agent alive in period t; the credit market clearing condition is hence $k_t = b_t$.

2.1 First-Best Allocation

The first best allocation is the allocation that would arise if no one in the economy had self-control problems. It is given by the solution to a fictitious social planner's consumption-saving problem where the planner has discounting with δ. The following Euler Equations characterize the first-best allocation, which we denote with superscript star throughout the paper:[11]

$$\begin{aligned} u'(c_i^*) &= \delta f'(k_i^*) u'(c_{i+1}^*), \text{ for } i = 0, 1, 2, .., I-1, \\ & \quad \text{and} \\ u'(c_I^*) &= \delta f'(k_I^*) u'(c_0'^*), \\ & \quad .. \end{aligned} \tag{1}$$

2.2 Implementing First-Best Allocation

Since people in this economy face self-control problems, laissez-faire market equilibrium cannot attain the first-best allocation. Our main interest in this paper is to find and characterize a tax system that implements the first-best 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 first-best allocation. However, since we focus our attention to linear equilibria, a linear tax system that implements the first-best allocation exists. We will verify this claim in Section 4.

2.3 Markov Equilibrim 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 at the steady-state are given by

$$\begin{aligned} R &= f'(k), \\ w &= f(k) - f'(k)k. \end{aligned} \tag{2}$$

[11] We do not state the transversality condition but our first-best allocation will converge to a steady state with positive capital as long as $k_0 > 0$.

In such a world, the only index we need to carry around is the age index i. In Appendix B, 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 first-best taxes do not depend on whether the economy is at the steady-state or in a transition.

Let τ_i be the savings (capital) tax agent $i = 0, 1, .., I$ 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 $\tau = \{\tau_i, T_i\}_i$. For each set of taxes, we define the policy functions $b_i(\cdot; \tau)$ for $i = 0, 1, .., I$, describing the optimal behavior of agent i given prices, taxes, and his beliefs about other agents' policy rules. 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 $\tau := \{\tau_i, T_i\}_{i=0}^{I}$ 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); (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.*

Parent's 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] \\
&\quad \text{s.t. for all } b_0 \qquad (3)\\
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\left[\pi_1 u\left(R\left(1-\tau_1\right)\hat{b}_1 + w + T_1 - b_2(\hat{b}_1)\right) + \left(1-\pi_1\right)W_1\left(\hat{b}_1;\tau\right)\right] \\
&\quad + \delta\beta_1\pi_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\} \\
&\quad \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\left[\pi_2 u\left(R\left(1-\tau_2\right)\hat{b}_2 + w + T_2 - b_3(\hat{b}_2)\right) + \left(1-\pi_2\right)W_2\left(\hat{b}_2;\tau\right)\right] \\
&\quad + \delta\beta_2\pi_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\} \\
&\quad \text{s.t. for all } b_2 \\
&\quad ... \\
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) + \\
&\quad + \delta\beta_{I-1}\left[\pi_{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\} + \left(1-\pi_{I-1}\right)W_{I-1}\left(\hat{b}_{I-1};\tau\right)\right] \qquad (4)\\
&\quad \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\left[\pi_I V\left(\hat{b}_I;\tau\right) + \left(1-\pi_I\right)W_I\left(\hat{b}_I;\tau\right)\right] \qquad (5)
\end{aligned}$$

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

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

with

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

Here, $V(b;\tau)$ represents the value of a parent's problem who saved b units 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 $b_1, ..., b_I$. 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 is sophisticated in the sense that he foresees this misalignment of preferences, and correctly forecasts future policies.

To understand the nested nature of policies better and the way we model partial sophistication, let us analyze the definition of policies in (4) and (5). First, constraint (5) 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. $(\beta_1, ..., \beta_I) \neq (1, ..., 1)$, 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. $(\beta_1, ..., \beta_I) = (1, ..., 1)$. 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.[12] Second, consider constraint (4) 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 (5), 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, (5), 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, (5), because the parent and self $I-1's$ sophisticated belief agree about how self I will behave.[13] 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.

[12] 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.

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

We restrict attention to linear equilibria, meaning equilibria with policy functions that are linear in net present value of current wealth. This implies that agents' problems are strictly concave maximization problems. As a result, first-order optimality conditions are not only necessary but also sufficient, which means we can replace agents maximization problems with the associated first-order conditions. First define

$$\begin{aligned}\Gamma_i(b) &= R(1-\tau_i)b + w + T_i + G_i,\\ G_i &= \frac{T_{i+1}}{R\,(1-\tau_{i+1})} + \frac{T_{i+2}}{R^2\,(1-\tau_{i+1})\,(1-\tau_{i+2})} + .. + \frac{T_I}{R^{I-i}\prod_{j=i+1}^{I}(1-\tau_j)} + \frac{T_0}{R^{I-i+1}(1-\tau_0)\prod_{j=i+1}^{I}(1-\tau_j)} + ...,\end{aligned}$$

where G_i is the net present value of future lump-sum taxes, $\Gamma_i(b)$ is the net present value of wealth available to an agent at the beginning of age $i+1$ with asset level b. We derive closed form solutions of the form:

$$c_i(b) = M_i \Gamma_{i-1}(b),$$

where the constant M_i is the fraction consumed out of net present value of wealth at the beginning of age i. The closed form is obtained by rewriting the parent's problem using linearity of the policy functions and the first-order approach, and finding analytic expressions for the value functions W_i and V and the vector of constants M_i describing the optimal linear policies.

Before we analyze the properties of optimal taxes in our environment however, we first want to analyze the main forces behind them in a simple three period model. This is the aim of the next section.

3 A three period example

In this section, we consider a simple heuristic example. The simplest environment to analyze the main mechanism must consist of three periods since we want to analyze the first-best tax on an agent, self 1, choosing a savings level, b_1, taking into account the action of a future self, self 2 choosing b_2, and future government policies, τ_2. To ease the exposition, we assume people initially have k_0 units of capital and production function has the form $f(k) = Rk$.

The problem of self 1 is:

$$\begin{aligned}\max_{b_1} u(k_0 - b_1) + \beta_1\delta\{\pi_1\,[u(R(1-\tau_1)b_1 + T_1 - b_2(b_1)) + \delta u(R(1-\tau_2)b_2(b_1) + T_2)]\\ +(1-\pi_1)\left[u(R(1-\tau_1)b_1 + T_1 - \hat{b}_2(b_1)) + \delta u(R(1-\tau_2)\hat{b}_2(b_1) + T_2)\right]\}\end{aligned}$$

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

$$b_2(b_1) \in \arg\max_{\tilde{b}_2} u(R(1-\tau_1)b_1 + T_1 - \tilde{b}_2) + \beta_2\delta u(R(1-\tau_2)\tilde{b}_2 + T_2), \tag{6}$$

$$\hat{b}_2(b_1) \in \arg\max_{\tilde{b}_2} u(R(1-\tau_1)b_1 + T_1 - \tilde{b}_2) + \delta u(R(1-\tau_2)\tilde{b}_2 + T_2), \tag{7}$$

where b_1, b_2 are self 1 and self 2's savings and $\hat{b}_2$ represents what self 1 naively believes self 2 will choose. The functions $b_2(\cdot)$ and $\hat{b}_2(\cdot)$ describe self 2's actual choice and naive self 1's expectation of self 2's choice as functions of first period savings, respectively. We plug these in the objective function of the parent.

For ease of exposition, assume the policies are differentiable.[14] The first-order conditions of the partially sophisticated self 1 is:

$$\begin{aligned} u'(c_1) &= \pi_1\delta\left[u'(c_2)[R(1-\tau_1)-b_2'(b_1)]+\delta R(1-\tau_2)b_2'(b_1)u'(c_3)\right] \\ &\quad +(1-\pi_1)\delta\left[u'(\hat{c}_2)[R(1-\tau_1)-\hat{b}_2'(b_1)]+\delta R(1-\tau_2)\hat{b}_2'(b_1)u'(\hat{c}_3)\right], \end{aligned}$$

where $\hat{c}_2, \hat{c}_3$ represent self 1's naive belief about self 2's consumption choice in periods two and three. First-order optimality conditions of (6) and (7) are:

$$\begin{aligned} u'(c_2) &= \beta_2\delta R(1-\tau_2)u'(c_3), \\ u'(\hat{c}_2) &= \delta R(1-\tau_2)u'(\hat{c}_3). \end{aligned}$$

The first condition above describes the actual behavior of self 2, and implies that self 2 should receive a subsidy, $\tau_2 = \frac{1}{\beta_2} - 1$. Since in this example self 2 does not face any future selves with self-control problems, the only role of period two tax is to correct for self 2's undersaving behavior, and that is why the sign of the optimal tax on self 2 is unambiguously negative. Once we plug these constraints in self 1's first-order optimality condition, we get:

$$u'(c_1) = \beta_1\delta R(1-\tau_1)u'(c_2)\left\{\pi_1\left[1+b_2'(b_1)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1)}\right]+(1-\pi_1)\frac{u'(\hat{c}_2)}{u'(c_2)}\right\}.$$

This means that optimal period one tax solves:

$$(1-\tau_1^*) = \frac{1}{\beta_1}\left\{\pi_1\left[1+b_2'(b_1^*)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1^*)}\right]+(1-\pi_1)\frac{u'(\hat{c}_2)}{u'(c_2^*)}\right\}^{-1},$$

where b_1^* and c_2^* represent first-best asset and consumption levels respectively. The tax formula for $1-\tau_1^*$ consists of two main components. The first part, $\frac{1}{\beta_1}$, is easier to understand. Because of his *current* self-control problem, self 1 discounts tomorrow by an extra β_1 and hence wants to undersave relative to the first-best. By multiplying the after tax return with $\frac{1}{\beta_1}$, we can exactly offset the extra discounting, thereby getting rid of this undersaving motive of the agent. Let us call this first part of the tax formula the *current component*. Clearly, the current component of tax is always negative, i.e. it always calls for a subsidy. It is intuitive that current component is not affected by the agent's sophistication level at all. This is not the end of the story, however. Self 1's choice of current savings is also affected by the actions of future selves and future government policies. Therefore, even if we correct for his undersaving through the current component of the tax, he still deviates from the first-best saving level in order to compensate for his future self's suboptimal

[14] It is well-known that in general we cannot guarantee even the continuity of the policy functions (e.g., see Morris and Postlewaite (1997), Krusell and Smith (2003), and Harris and Laibson (2000)).

actions (due to future self-control problems) and/or in response to future policies. The second part of the tax formula $\left\{\pi_1\left[1+b_2'(b_1)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1^*)}\right]+(1-\pi_1)\frac{u'(\hat{c}_2)}{u'(c_2)}\right\}^{-1}$ is there to correct deviations in current savings caused by future actions and policies. We call this part the *future component* of the tax formula. This is where the level of sophistication matters. As we show in Appendix A, the future component is always less than one, i.e. *it calls for a tax, independent of the level of sophistication.*[15]

To gain intuition on why future component always calls for a tax, first consider the future component in the case where self 1 is fully sophisticated, i.e. set $\pi_1 = 1$. For the future component to be less than one under full sophistication, it must be that $b_2'(b_1^*)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1^*)} > 1$; a property always satisfied since $-1+\frac{1}{\beta_2} > 0$ and $b_2'(b_1^*) > 0$. The later condition (monotonicity of the policy) is quite general. The condition $-1+\frac{1}{\beta_2} > 0$ is key, and reflects the fact that from self 1's perspective self 2 undersaves in period two. This is a violation of an Envelope condition that holds when agents are time-consistent. In the quasi-hyperbolic model with sophisticated agents, according to self 1, self 2 is undersaving, and this appears as an extra return to savings for self 1. Since the allocation satisfies the condition $u'(c_2^*) = \beta_2\delta(1-\tau_2^*)Ru'(c_3^*)$, each unit saved by self 2 has a cost $u'(c_2^*)$ that is lower than the self 1's perceived return $\delta(1-\tau_2^*)Ru'(c_3^*)$. Hence, whenever $b_2'(b_1^*) > 0$, self 1 is induced to save some extra money. The reader might still feel puzzled by our argument: after all, when facing the optimal taxes, self 2 saves the right amount with respect to first-best. Note however, that from self 1's perspective, self 2 is still undersaving (at the new price that is inflated by the subsidy).

Now consider the problem of a self 1 who is fully naive, that is, set $\pi_1 = 0$. The key condition for a tax is now $\frac{u'(\hat{c}_2)}{u'(c_2^*)} > 1$. Since self 1 naively believes that self 2 has no self-control problems, he thinks self 2 will save more than what self 2 will actually save. More precisely, self 1 incorrectly believes that period two consumption $\hat{c}_2$ will be low compared to c_2^*, and hence, marginal utility of consumption in period two is higher relative to what it really is. As a result, naive self 1 perceives a large return $Ru'(\hat{c}_2)$ to his savings and saves too much.

We have just seen that both fully sophisticated and fully naive agents have a tendency to oversave to compensate for their future selves' behavior. Since under both full sophistication and full naivete future component calls for a tax, it calls for a tax under any level of partial sophistication as well.

The sign of the optimal capital tax depends on whether the current or the future component dominates. For $\beta_i \equiv \beta < 1$ (i.e., constant self-control) the current component always dominates, and hence, the optimal tax is negative (i.e., optimality calls for a capital subsidy). We will see below that for β_i changing with age the optimal tax can in general be positive or negative.

[15] However, as we show in Appendix D, the *magnitude* of the future component, and therefore, the level of optimal taxes in general depend on the level of sophistication.

4 Optimal Taxes

In this section we analyze optimal capital taxes in the model introduced in section 2. The first proposition below characterizes optimal taxes when utility is logarithmic for any level of sophistication.

Proposition 1 *Suppose $u(c) = \log(c)$. For any level of partial sophistication over the life-cycle, $\pi = (\pi_1, \pi_2, .., \pi_I)$, we have:*

$$1 - \tau_0^* = 1 - \delta + \beta_1 \delta,$$
$$1 - \tau_i^* = \frac{1}{\beta_i} \left(1 - \delta + \beta_{i+1}\delta\right) \text{, for } i \in \{1, .., I-1\}$$
$$1 - \tau_I^* = \frac{1}{\beta_I}.$$

Proof. Relegated to Appendix B. ■

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. Since our model of partial sophistication is different from that considered in the literature, it is interesting that it shares this property with the more standard framework. In Appendix A, we use the 3-period example of Section 3 to provide further intuition on this equivalence result.

Proposition 1 holds regardless of whether the economy is in the steady state or in a transition. Below we show that if the economy is in the steady state and all the agents in the economy are fully sophisticated, then optimal taxes characterized above for the $\sigma = 1$ case is valid for any σ.

Proposition 2 *Assume k is such that $\delta f'(k) = 1$ and $\pi_i = 1$ for all i, then optimal taxes are independent of CEIS coefficient σ.*

Proof. Relegated to Appendix B. ■

Using the jargon developed in Section 3, $\frac{1}{\beta_i}$ is the current component of the optimal tax on agent at age i and $\left(1 - \delta + \beta_{i+1}\delta\right)$ is the future component. As expected, the current component always calls for a subsidy whereas the future component always calls for a tax in order to implement first-best allocation. Since τ_0^* is only shaped by the future component, it is 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 analyze taxation of wealth transferred across generations. We study this topic in detail in Pavoni and Yazici (2012).

4.1 Lessons for Capital Taxation

Proposition 1 implies several general lessons for capital taxes which are summarized below in a series of corollaries.

Corollary 3 ***(Age-dependence)*** *Optimal capital taxes are age-dependent.*

The first lesson to be learnt from logarithmic utility case is that in general optimal capital taxes should depend on people's age. The reason for the necessity of this dependence is the changing the degree of self-control problem over age, for which, as discussed in the introduction, there is an overwhelming amount of evidence in personality psychology literature.

Corollary 4 ***(Sign of the Capital Taxes)***
(1) Optimal capital taxes might be positive or negative depending on how β_i changes with i.
(2) With log utility, if $\beta_{i+1} \geq \beta_i$, for all i, optimal capital tax is negative for all ages:

Proof. (1) For an example of $\tau_i > 0$, set $\beta_{i+1} \approx 0$ and $\beta_i > 1 - \delta$. For a subsidy, set $\beta_i = \beta_{i+1} = \beta < 1$. See also Figure 1.
(2) $1 - \tau_i^* = \frac{1}{\beta_i}(1 - \delta + \beta_{i+1}\delta) > \frac{\beta_{i+1}}{\beta_i} \geq 1$. ■

The lesson to be taken about the sign of the age-dependent capital taxes is simply that optimal capital taxes might be positive or negative depending on the evolution of the severity of the self-control problem over the life cycle. This is an important message since it 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 is quite contrary to the presumption in the literature that self-control problems always imply subsidies.[16] The existing literature overlooks this result because they assume constant self-control problems under which the current component always dominates the future component, and hence, implementing first-best calls for subsidies.

There is also a sharper message when we take logarithmic utility or CEIS with full sophistication seriously: if, as suggested by personality psychologists, the degree of self-control problem is decreasing over the life-cycle, then capital should be subsidized at all ages.

Corollary 5 ***(Monotonicity of Capital Taxes)***
(1) Optimal capital taxes might be increasing or decreasing depending on how β_i change with i.
(2) With log utility, optimal capital taxes are always negative in the last period before parenthood, $\tau_I^* < 0$. If $\beta_{i+1} \geq \beta_i$, for all i, then optimal capital tax is negative for all ages.

Proof. (1) See the green line with crosses in Figure 2 below for an example.
(2) $1 - \tau_{i-1}^* = \frac{1-\delta}{\beta_{i-1}} + \frac{\beta_i\delta}{\beta_{i-1}} > \frac{1-\delta}{\beta_i} + \frac{\beta_i\delta}{\beta_{i-1}} > \frac{1-\delta}{\beta_i} + \frac{\beta_{i+1}\delta}{\beta_i} = 1 - \tau_i^*$, 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 lesson about the monotonicity of capital taxes again points to the importance of the evolution of the severity of self-control problem over the life-cycle: without the knowledge of how β_i changes with i, policy prescriptions regarding the optimal dependence of capital taxes on age would be misguided. Point (2) in Corollary 5 is potentially important since there is evidence in the personal psychology literature which suggests that people's ability to self-control increases

[16]O'Donoghue and Rabin (1999) is an exception where it says if the agent is sophisticated then he may oversave. However, even in that paper it says that "naifs will undersave in essentially any savings model" and hence should be subsidized. Proposition 1 shows that, in our environment depending on how self-control evolves over the life-cycle, even naive may oversave and hence may need to be taxed.

concavely with age.[17] Interestingly, this result is contrary to Krusell, Kuruscu, and Smith (2010) which concludes unambiguously that in any finite economy with constant self-control, subsidies should be increasing with age.

We display Figure 1 and Figure 2 to show how different assumptions about the pattern of self-control problem over the life-span can affect the evolution of optimal capital taxes. In the first figure, we see that constant β_i, which is depicted by a dashed line, implies constant subsidies as found by previous literature. The decreasing pattern of β_i depicted by the red crosses on the left panel of the figure delivers capital taxes to be positive until the very last period as shown on the right panel. In the second figure, we see different self-control patterns that are all increasing with age. In this case, as the theory shows, capital should be subsidized; however, we see that the monotonicity property of subsidies with respect to age depends on the curvature of β_i.

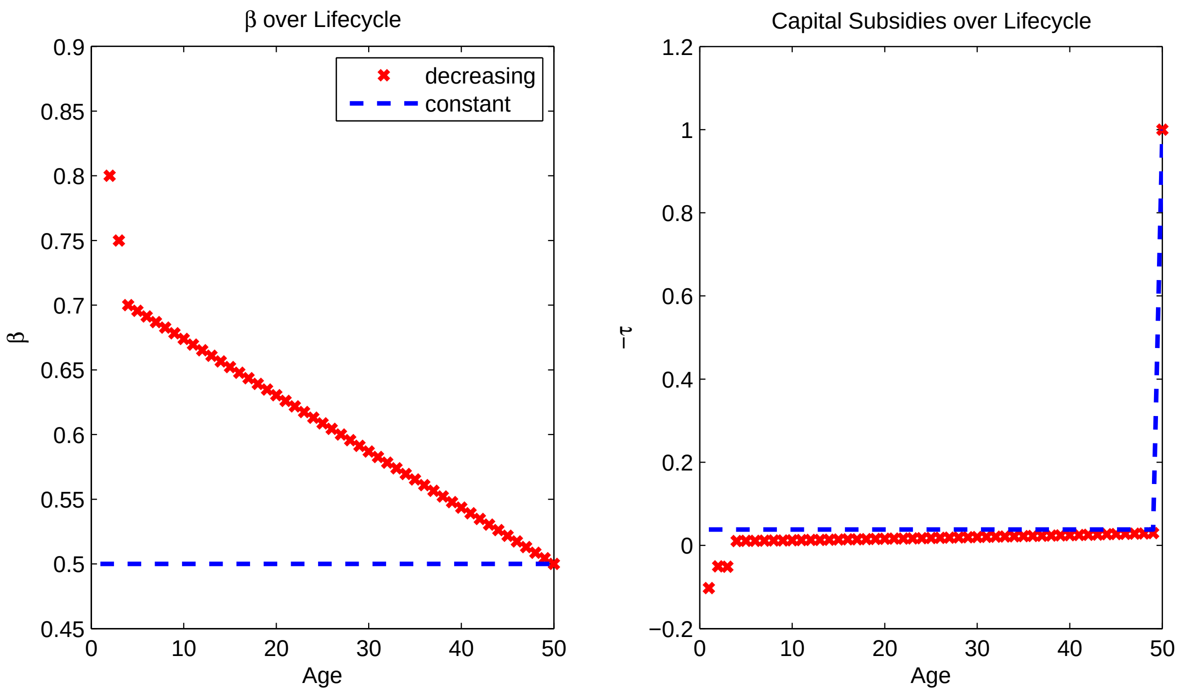

Figure 1: Optimal capital subsidies for decreasing and constant patterns of β_i over the life-cycle.

4.2 Numerical Analysis

In this subsection, we numerically analyze 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 steady-state condition holds and all the agents in the model are fully sophisticated. In order to conduct a numerical analysis, we have to choose particular values

[17] See John, Gosling, Potter, and Srivastava (2003) and Roberts, Walton, and Viechtbauer (2006), among others, and the next section for more details on the evidence on concavity of β_i as a function of i.

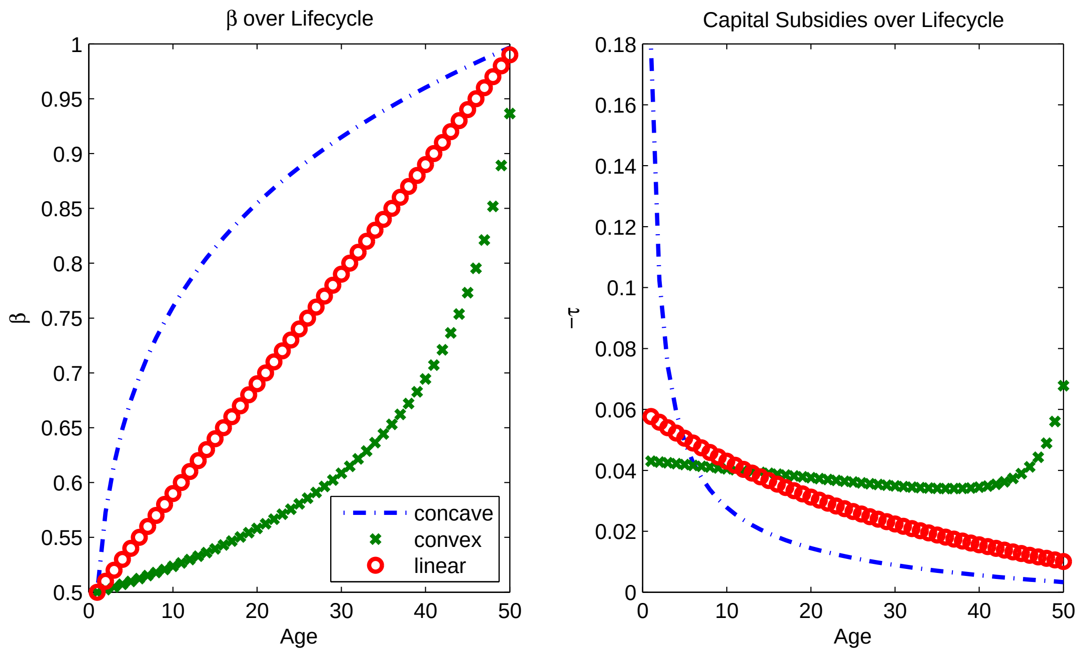

Figure 2: Optimal capital subsidies for concave, linear, and convex increasing patterns of β_i over the life-cycle.

for the parameters of the model. Individuals are assumed to be born at the real-time age of 20 and they live $I = 50$ years, so they die at age 70. 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. So, we do not specify values for these parameters. The only parameters that are needed are the true yearly discount factor δ and the evolution of self-control parameter with age, $\{\beta_i\}_i$.[18] We set the true yearly discount factor $\delta = 0.96$ which is consistent with Laibson, Repetto, and Tobacman (2007)'s estimate in a constant self-control model when $\sigma = 1$.

As evident from the optimal tax formulas, self-control function $\{\beta_i\}$ is the crucial 'parameter'. Moreover, Figure 1 and Figure 2 show that taxes are in general very sensitive to the vector $\{\beta_i\}$. Therefore, to say something concrete, we need to make several assumptions about $\{\beta_i\}$. We assume that β_i is increasing and concave in i. In words, this means that the degree of self-control problem decreases with age and this decline slows down with age. We have two sets of evidence in favor of these assumptions. First, research on intertemporal 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.[19] Second, personality

[18]Observe that if we want the taxes computed using the formulas in Proposition 1 to be valid under any σ and with full sophistication, then we need to assure that the interest rate R (or the deeper parameters of the production function F and d) satisfies the steady-state condition

$$R = f'(k) = F'(k, 1) - d = \delta^{-1},$$

where k refers to steady-state level of capital stock.

[19]See Green, Fry, and Myerson (1994) , Green, Myerson, and Ostaszewski (1999), Read and Read (2004), and Ameriks, Caplin, Leahy, and Tyler

psychologists associate self-control with conscientiousness, one of the 'big five' personality factors,[20] and in the words of Caspi, Roberts, Robins, and Trzesniewski (2003) 'it appears that the increase in conscientiousness is one of the most robust patterns in personality development, especially in young adulthood.' So, there seems to be a consensus among psychologists that self-control increases with age. The evidence for concavity of this increase comes again from the personality psychology literature. John, Gosling, Potter, and Srivastava (2003) and Roberts, Walton, and Viechtbauer (2006) both find that conscientiousness increases concavely over the life-cycle. John, Gosling, Potter, and Srivastava (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.'[21] We use a quadratic β_i function and perform robustness checks by varying the degree of concavity allowing for linearity as well.

We also make assumptions about the level of β_i at the youngest and oldest age. In our benchmark calculations, we assume $\beta_1 = 0.5$. For $\sigma = 1$, Laibson, Repetto, and Tobacman (2007) estimate $\beta = 0.81$ along with $\delta = 0.96$. However, in their sample, the average age is around 40. Using the linear quadratic developmental trend suggested by personality psychology, we compute that self-control parameter at age 20, β_1, should be approximately 0.5 in order to have this parameter equal to 0.81 at age 40. We check for robustness by setting $\beta_1 = 0.4$ and $\beta_1 = 0.65$.[22] We assume that self-control problem vanishes towards the end of one's life-cycle. This is in line with the evidence from research on intertemporal discounting as summarized in Read and Read (2004): "Green et al's major result- that younger people show hyperbolic discounting while older people show exponential discounting - is supported by our data." The old people have a mean age of seventy-five in Read and Read (2004) and seventy in Green, Myerson, and Ostaszewski (1999), which is consistent with the age of our oldest agent, seventy.

Now we report the results. In all our simulations, capital taxes are negative so they are indeed subsidies and these subsidies are decreasing with age throughout the life-cycle. This is expected given Corollary 4 and Corollary 5 and the fact that we followed empirical evidence by assuming β_i is increasing and concave in i. As Figure 3 and Figure 4 depict, optimal capital subsidies at the beginning of the life-cycle are between 4% and 12%, depending on the specifications, and decrease monotonically with age to between 0 and 1%.[23] The speed at which the subsidies decline depends on the curvature of β_i. Figure 3 displays the sensitivity of subsidies to the degree of concavity of β_i function. In the figure, the β_i function is depicted on the left while the corresponding age-dependent capital subsidies are depicted on the right. The

(2007).

[20] Ameriks, Caplin, Leahy, and Tyler (2007) also analyzes the 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.' Borghans, Duckworth, Heckman, and ter Weel (2008) also states that conscientiousness is conceptually related to self-control problems.

[21] It is possible to compute one-year short-term discount rates (our $\beta's$) using Green, Myerson, and Ostaszewski (1999)'s estimates of hyperbolic discount functions for different age groups in his study and such an analysis confirms that β is a concave increasing function of age. However, he has only three age groups.

[22] 0.65 is the one-year short-term discount rate we computed using Green, Myerson, and Ostaszewski (1999)'s estimate of hyperbolic discount function for his young adult group which has mean age of 20 years.

[23] If we denote by β_{51} the self-control parameter of the parent (of course, also denotable as β_0), then - by construction - $\beta_{51} = 1$. Although it is not always easy to see in the figures, in our specifications however, β_{50} is typically less than one.

initial level of self-control problem β_1 is set to 0.5. The three curves other than the solid one represent different degrees of concavity within the same family of quadratic functions. The blue dashed curve has the highest level of concavity whereas the green straight line has the least, and the red dotted curve is in between.[24] We see that the more concave the function is the higher the initial level of taxes are, the faster the decline with age is, and the lower the value of final level of taxes. In the solid curve in turquoise, β_i is a 4th root function of age. The corresponding age-dependent capital subsidies curve on the right shows that the type of concave function chosen also matters for the level and shape of age-dependent subsidies. Figure 4 shows sensitivity with respect to our assumption regerding the initial level of the self-control problem β_1. In the figure, we assume β_i follows a quadratic pattern with coefficient $b = -.02$, as the blue dashed line on Figure 3. We see that the lower is β_1 the higher is the starting value of taxes and also the sharper the decline.

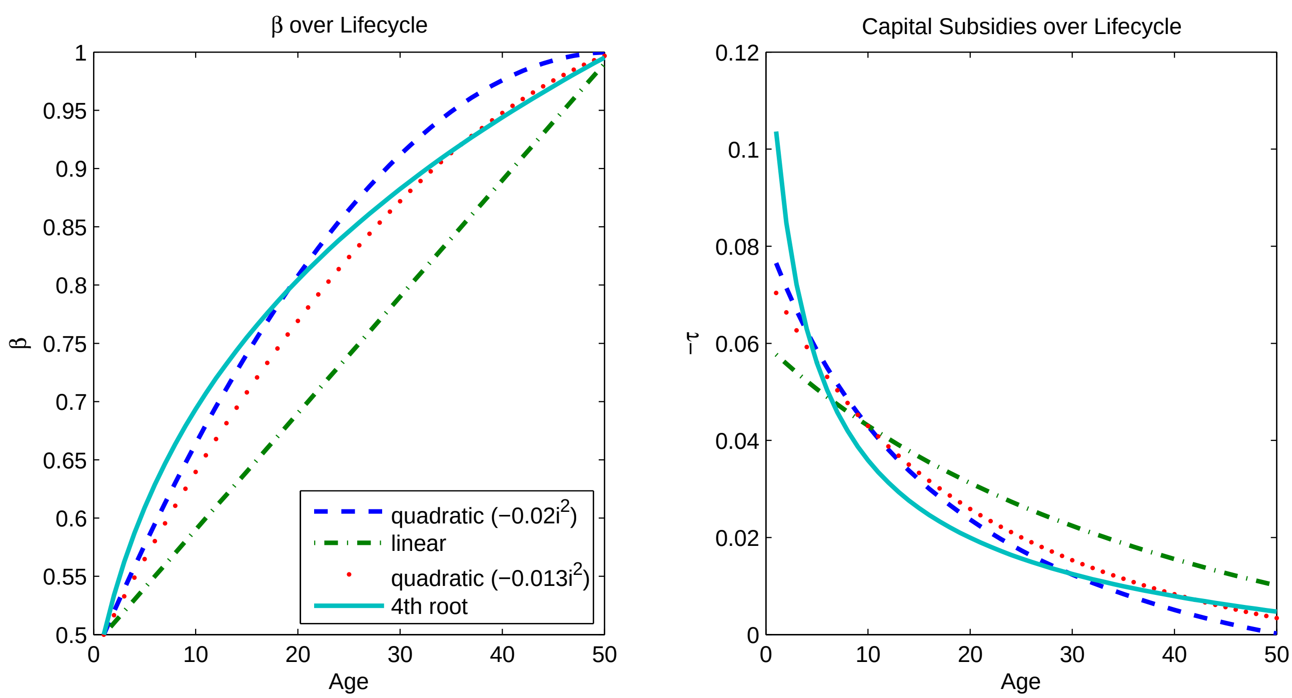

Figure 3: Sensitivity with respect to the curvature of β_i

Observe that here the tax base is the gross return on asset holdings. Most actual tax systems, however, tax *asset income*. If we translate our numbers taking that into account, we find that optimal subsidies on capital income at early ages always takes values above 100%, touching the level of almost 300% for the highest range of the subsidy.[25] These are

[24]The exact quadratic form used is $\beta_i = ai + bi^2 + d$. On the picture we only report the coefficient of the quadratic term, b, since the remaining parameters a and d are pinned down by the initial and final values of the self-control parameter, $\beta_1 = .5$ and $\beta_{51} = 1$.

[25]Denoting the capital income tax by τ_i^k, the relation between our taxes and tax on capital income is: $1 - \tau_i^k = \frac{R(1-\tau_i)-1}{R-1}$. As a consequence, for

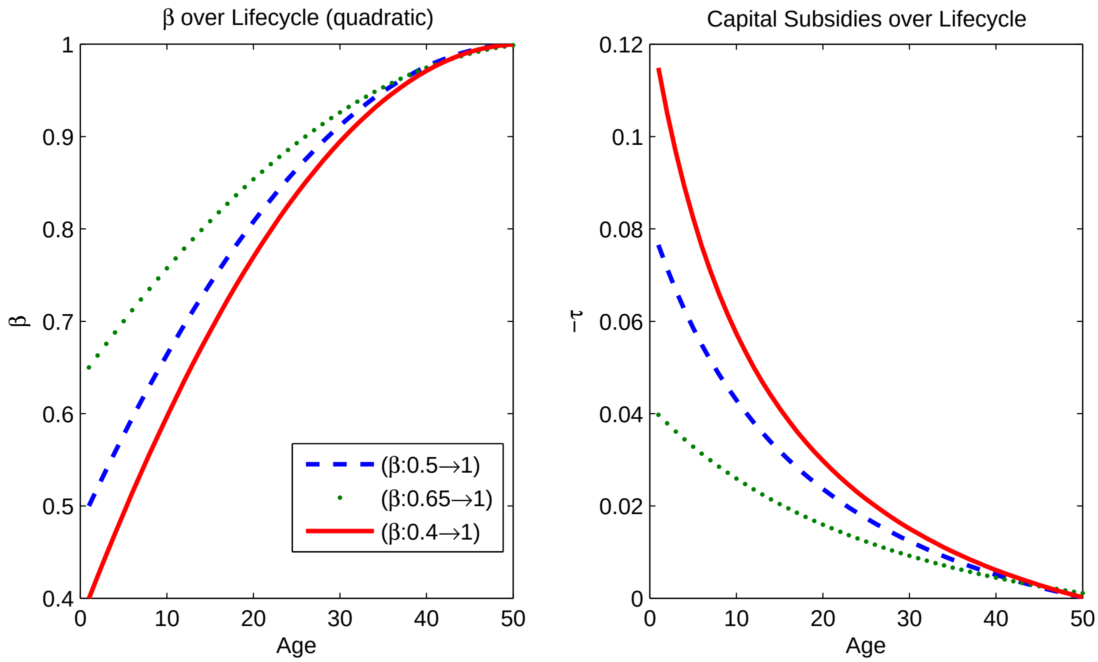

Figure 4: Sensitivity with respect to β_1

Table 1: Optimal Subsidies under Constant Self-Control Problems

β	0.81	0.65	0.5	0.4
$-\tau$	0.93%	2.15%	4%	6%
$-\tau^k$	23%	54%	100%	149%

obviously large numbers.

Now we argue that the subsidies we obtain are much larger than those implied by models with constant self-control, and explain why. Table 1 reports the taxes implied by constant self-control for the four representative levels of β. There are two main forces that inflate the optimal level of subsidies when β_i changes over the life-cycle. First, the empirical evidence implies quite low β's for young individuals. In the jargon of Section 3, the *current component* of the tax is quite strong for young individuals. Second, the increasing level of self-control with age reduces the importance of *the future component* of the tax. Thus, for empirically plausible patterns of self-control, both forces go in the direction of increasing savings subsidy to the young.

Interestingly, the mitigating effect of increasing self-control on the future component of optimal taxes is strong enough to imply larger taxes for older individuals as well, compared to the model with constant self-control. Looking, for

$-\tau_1 = .12$ we have $-\tau_1^k = 2.997$.

example, at the column related to a constant $\beta = .81$ in Table 1, we see that the implied optimal capital income subsidy $-\tau^k$ is 23%. In our model, the agent with the same degree of self-control problem receives capital income subsidy of more than 57%. To see this, observe that the blue dashed line in the left panel of Figure 3 shows that in our model an agent has the corresponding degree of self-control in period twenty, i.e., $\beta_{20} \approx .81$, and the corresponding line on the right panel shows $-\tau_{20} \approx .023$, which indeed implies a capital income tax larger than 57%. Another way of noticing the same phenomenon is to realize that in all our parameterizations, for the oldest individuals we have β_{50} well above .9. In spite of that, capital subsidies are sometimes larger than those implies by a model with constant self-control with $\beta = .81$.

5 Effect of Partial Sophistication

In the previous section, we show that: (1) when the constant relative risk aversion coefficient σ is equal to 1, then the degree of sophistication is immaterial for taxes; (2) under the assumption that all the agents in the economy are fully sophisticated and the economy is at a steady-state, σ is immaterial for taxes. In these two cases, taxes are given by Proposition 1. It is evident that in order to investigate the robustness of our policy findings with respect to naivete, we need to move away from the assumptions of $\sigma = 1$ and full sophistication at the same time. This is exactly what this section does. Unfortunately, when $\sigma \neq 1$ and agents are allowed to be partially sophisticated, we do not get closed form solutions for optimal taxes. Therefore, we have to resort to numerical analysis. For simplicity, we keep the assumption that the economy is at steady state.

In our first set of analysis, we set $\sigma = 2$ and analyze how different patterns of the evolution of partial sophistication over the life-cycle affect optimal life-cycle subsidies.[26] In Figure 5, the blue solid curve represents the benchmark case of full sophistication ($\pi = 1$) where optimal taxes do not depend on σ. 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 repeat this numerical analysis for $\pi = 0.3, 0.5, 0.7$, and 0.9. As evident from Figure 5, 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. For instance, for $\pi = 0.3$, the level of sophistication in period 9 is 0.3 whereas it is 1 in period 10. As a result, as the figure shows, optimal period 9 subsidies are significantly larger compared to the benchmark case. Similarly, a significant decline in sophistication from 1 in period 10 to 0.3 in period 11 implies much lower optimal subsidies compared to the benchmark case. On the other hand, since sophistication level is 0.3 in both periods 11 and 12, optimal subsidies in period 11 are roughly identical to the case of fully sophisticated benchmark. In other words, when the level of sophistication does not change across periods, its level is not quantitatively important for the level of optimal taxes. We also analyze the effect of sophistication on optimal

[26]We set $\beta_1 = 0.5$ and $\beta_{51} = 1$, and $\beta_i = ai + bi^2 + d$, where $b = -0.02$ throughout this section.

subsidies when the level of sophistication changes smoothly over the lifecycle. We assume π increases concavely. This experiment is summarized in Figure 6, where we confirm that the level of sophistication matters for optimal taxes only when it changes sharply between two adjacent periods.

Finally, we do robustness checks for σ different from 2. As Figure 6 suggests, as σ moves away from 1, the effect of sophistication becomes more significant. However, even when $\sigma = 5$, the difference between optimal capital subsidies in the benchmark model 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. Figure 6 suggests a qualitative pattern regarding how optimal taxes are affected by sophistication level for a given level of σ. When $\sigma = 0.5$, the optimal subsidies under partial sophistication are given by the dotted line that lies below the solid curve, which also represents optimal taxes for $\sigma = 0.5$ under full sophistication. On the other hand, for all $\sigma > 1$ in the figure, we see that optimal subsidies under partial sophistication are higher than optimal subsidies under full sophistication at every age level. These observations suggest a particular pattern: that for $\sigma > 1 (< 1)$, optimal taxes increase with the level of sophistication. In Appendix D, we explain this pattern formally.[27]

So, there are two major conclusions derived from the above set of experiments. First, as long as the level of naivete is not changing abruptly from one period to another, the level optimal capital subsidies over the lifecycle is robust to various scenarios about how sophistication changes with age. Second, as the last experiment shows, when the level of partial sophistication is changing smoothly (or not changing at all), the level optimal capital subsidies over the lifecycle is not significantly affected by our choice of the coefficient of constant relative risk aversion.

6 Conclusion and Discussion

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 assume that the severity of the self-control problem changes over the life-cycle. We also allow for age-dependent partial sophistication. 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, then capital should be subsidized and the subsidy should decrease with age. Our numerical analysis shows that capital subsidies should start somewhere between 4% and 12% at the beginning of the life-cycle and decline monotonically with age to somewhere between 0% and 1%, depending on the particular parameterization of the model. These are very large numbers, especially if we translate them into subsidies to capital income. More importantly, we show they are much larger than the savings subsidy we would obtain in models with constant self-control. This is

[27] 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 show that this result is valid under our way of modeling partial sophistication as well.

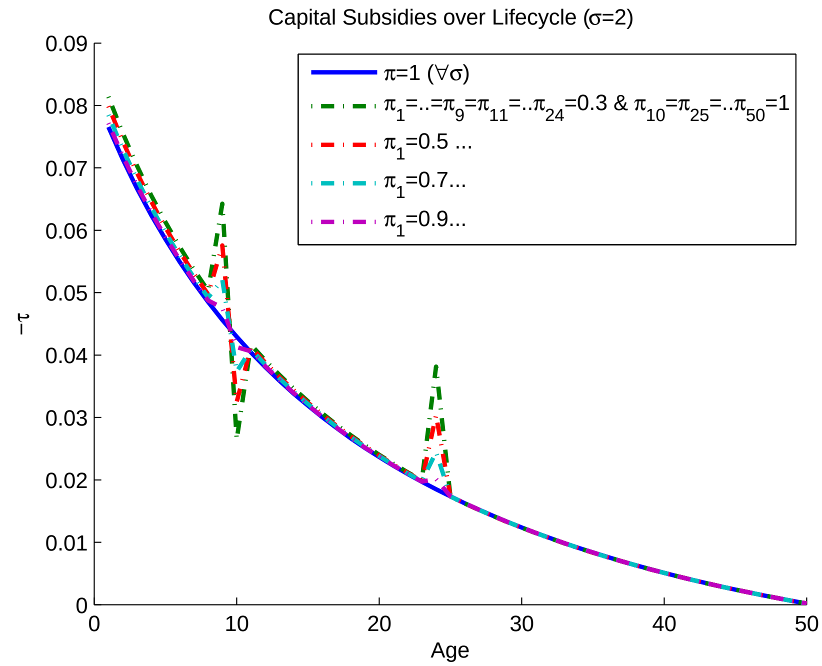

Figure 5: Sensitivity with respect to partial sophistication (jumps in π)

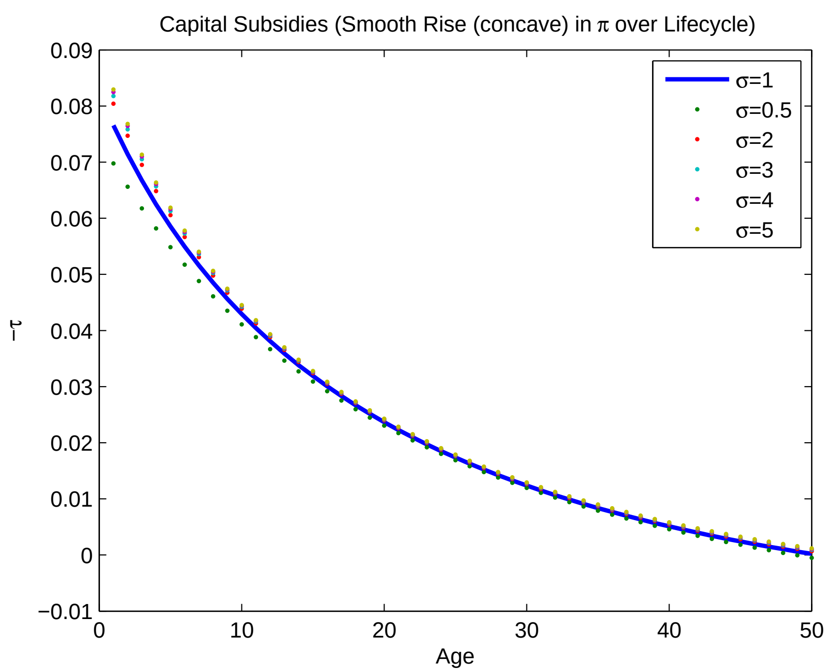

Figure 6: Sensitivity with respect to partial sophistication (smoothly rising π)

especially true for young individuals, but we explain why it holds actually at almost all ages. Our model is probably too simple for delivering precise policy predictions. We believe however, that our analysis provides a key lesson: researchers who take self-control problems seriously should also carefully measure the evolution of self-control problems over the life-cycle seriously before making policy suggestions.

We conclude by arguing that existence of illiquid assets does not change our optimal tax results as long as there are no borrowing constraints. More precisely, 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. Appendix E provides a more formal illustration of this argument through a three period example.

References

ALI, S. N. (2011): "Learning Self-Control," *The Quarterly Journal of Economics*, 126(2), 857–893.

AMERIKS, J., A. CAPLIN, J. LEAHY, AND T. TYLER (2007): "Measuring Self-Control Problems," *American Economic Review*, 97(3), 966–972.

ARIELY, D., AND K. WERTENBROCH (2002): "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science*, 13(3), 219–224.

BORGHANS, L., A. L. DUCKWORTH, J. J. HECKMAN, AND B. TER WEEL (2008): "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 43(4).

CASPI, A., B. W. ROBERTS, R. W. ROBINS, AND K. H. TRZESNIEWSKI (2003): *Handbook of Life Course*.

DELLAVIGNA, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47(2), 315–72.

EROSA, A., AND M. GERVAIS (2002): "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 105(2), 338–369.

FREDERICK, S., G. LOEWENSTEIN, AND T. O'DONOGHUE (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40(2), 351–401.

GREEN, L., A. F. FRY, AND J. MYERSON (1994): "Discounting of Delayed Rewards: A Life-Span Comparison," *Psychological Science*, 5, 33–36.

GREEN, L., J. MYERSON, AND P. OSTASZEWSKI (1999): "Discounting of Delayed Rewards across the Life-Span: Age Differences in Individual Discounting Functions," *Behavioural Processes*, 46, 89–96.

GUL, F., AND W. PESENDORFER (2001): "Temptation and Self-Control," *Econometrica*, 69(6), 1403–1435.

HARRIS, C., AND D. LAIBSON (2000): "Hyperbolic Discounting and Consumption," *8th World Congress of the Econometric Society*.

HELSON, R., C. JONES, AND V. S. Y. KWAN (2002): "Personality Change over 40 Years of Adulthood: Hierarchical Linear Modeling Analyses of Two Longitudinal Samples," *Journal of Personality and Social Psychology*, 83(3), 752–766.

JOHN, O. P., S. D. GOSLING, J. POTTER, AND S. SRIVASTAVA (2003): "Development of Personality in Early and Middle Adulthood: Set Like Plaster or Persistent Change?," *Journal of Personality and Social Psychology*, 84(5), 1041–1053.

KOCHERLAKOTA, N. (2010): *The New Dynamic Public Finance*, The Toulouse Lectures in Economics. Princeton University Press, ISBN: 9780691139159.

KRUSELL, P., B. KURUSCU, AND A. J. SMITH (2002): "Equilibrium Welfare and Government Policy with Quasi-geometric Discounting," *Journal of Economic Theory*, 105(1), 42–72.

——— (2010): "Temptation and Taxation," *Econometrica*, 78(6), 2063–2084.

KRUSELL, P., AND A. A. SMITH (2003): "Consumption–Savings Decisions with Quasi–Geometric Discounting," *Econometrica*, 71(1), 365–375.

LAIBSON, D. (1994): "Self-Control and Saving," Discussion paper.

——— (1997): "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112(2), 443–77.

LAIBSON, D., A. REPETTO, AND J. TOBACMAN (2007): "Estimating Discount Functions with Consumption Choices over the Lifecycle," NBER Working Papers 13314, National Bureau of Economic Research, Inc.

MORRIS, S., AND A. POSTLEWAITE (1997): "Observational Implications of Non-Exponential Discounting," Discussion paper.

O'DONOGHUE, T., AND M. RABIN (1999): "Doing It Now or Later," *American Economic Review*, 89(1), 103–124.

——— (2001): "Choice And Procrastination," *The Quarterly Journal of Economics*, 116(1), 121–160.

——— (2003): "Self Awareness and Self Control," in *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, pp. 217–243. Russell Sage Foundation Press.

PAVONI, N., AND H. YAZICI (2012): "Optimal Taxation of Bequests and Inter-Vivos Transfers," Working paper.

PHELPS, E. S., AND R. A. POLLAK (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *The Review of Economic Studies*, 35(2), pp. 185–199.

POLLAK, R. A. (1968): "Consistent Planning," *The Review of Economic Studies*, 35(2), 201–208.

READ, D., AND N. READ (2004): "Time Discounting over the Lifespan," *Organizational Behavior and Human Decision Processes*, 94(1), 22–32.

ROBERTS, B. W., K. E. WALTON, AND W. VIECHTBAUER (2006): "Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies," *Psychological Bulletin*, 132(1), 1–25.

STROTZ, R. H. (1955): "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies*, 23(3), 165–180.

THALER, R. H., AND H. M. SHEFRIN (1981): "An Economic Theory of Self-Control," *Journal of Political Economy*, 89(2), 392–406.

7 Appendix A - Explaining Future Component of Optimal Taxes

In this section we analyze the future component described in Section 3 more closely. To isolate the future component, we assume $\beta_1 = 1$. The analysis is heuristic in the sense that we assume that the equilibrium involves differentiable policies.

Sophisticated Future Component. We first analyze the future component when the agent is fully sophisticated in period one, so suppose $\pi_1 = 1$. We call the optimal tax on sophisticated period one agent τ_1^S. The first order condition of self 2 reads as follows:

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

This implies that in order to implement first best period two saving, b_2^*, the planner has to subsidize the return to period two saving by $(1-\tau_2^*) = \frac{1}{\beta_2}$.

Self1's first order conditions reads:

$$u'(c_1) = \delta \left[u'(c_2)R(1-\tau_1) + b_2'(b_1)\left\{ -u'(c_2) + \delta R(1-\tau_2)u'(c_3) \right\} \right].$$

The right-hand-side of this condition is the marginal benefit of saving an extra unit in period one whereas the left-hand-side is the marginal cost. When self 2 has no self-control problem, $u'(c_2) = \delta R(1-\tau_2)u'(c_3)$, so the right-hand side reduces to $u'(c_1) = \delta u'(c_2)R(1-\tau_1)$. However, being fully sophisticated, self 1 correctly believes that self 2 has a self-control problem and is going to undersave from self 1's perspective (from first-best perspective self 2 is saving the right amount thanks to period two subsidy), meaning, $u'(c_2) < \delta R(1-\tau_2)u'(c_3)$. Self 1 correctly believes that saving an extra unit in period one has an additional marginal benefit of increasing b_2, which is equal to $\delta b_2'(b_1)\left\{ -u'(c_2) + \delta R(1-\tau_2)u'(c_3) \right\} > 0$. As a result, he keeps increasing his savings until

$$\begin{aligned} u'(c_1) &= \delta \left[u'(c_2)R(1-\tau_1) + b_2'(b_1)\left\{ -u'(c_2) + \delta R(1-\tau_2)u'(c_3) \right\} \right] \\ &> \delta u'(c_2)R(1-\tau_1) \end{aligned}$$

which implies

$$(1-\tau_1^S) < \frac{u'(c_1^*)}{\delta R u'(c_2^*)} = 1,$$

meaning self 1 should be taxed for his oversaving. The exact amount of the tax solves:

$$1-\tau_1^S = \left[1 + b_2'(b_1^*)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1^S)}\right]^{-1} < 1$$

since $b_2'(b_1) > 0$.

It is important here to realize that even though self 2 saves the right amount with respect to first-best thanks to period two taxes, from self 1's perspective he is undersaving and that is why self 1 wants to oversave and hence should be taxed. So, even if self 2 was an oversaver and we had to tax him to make him save at the first best level, as long as $\beta_2 < 1$, the argument would still apply and we would still have to tax self 1. Moreover, if $\beta_2 = 1$, then even if we had to distort the problem of self 2, $\tau_2 \neq 0$, to make self 2 save at the first-best level, we would still have future component of self 1's tax equal to zero. So, what matters for future component is not future government policy but it is next period self's degree of self-control problem.

Naive Future Component. Now we analyze the future component when self 1 is fully naive, so suppose $\pi_1 = 0$. We call the optimal tax on the naive self 1 agent τ_1^N. As in the previous case where self-1 was sophisticated, self-2's first-order condition

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

implies in order to implement first best we have to subsidize self 2 by $(1-\tau_2^*) = \frac{1}{\beta_2}$.

The naive self 1 incorrectly believes that self 2 chooses his savings according to:

$$u'(\hat{c}_2) = \delta R(1-\tau_2)u'(\hat{c}_3),$$

meaning for any level of b_1 self 1's guess of self 2's consumption is less than self 2's actual consumption, $\hat{c}_2 < c_2$. As a result, without any period one tax, self 1 would incorrectly think that if he sets period one saving equal to b_1^* period two consumption would be too low since

$$u'(c_1^*) = \delta R u'(c_2^*) < \delta R u'(\hat{c}_2).$$

This implies that without period one tax self 1 would set his savings above b_1^* since self 1's first order condition for b_1 is:

$$u'(c_1) = \delta R(1-\tau_1^N)u'(\hat{c}_2).$$

So, to prevent this oversaving, we need to tax b_1, and the exact amount of the tax is given by:

$$(1-\tau_1^N) = \frac{u'(c_2^*)}{u'(\hat{c}_2)} < 1.$$

Again it is important to note that whether self 2 is an oversaver or an undersaver from first-best perspective does not matter for the result that self 1 is an oversaver and hence should be taxed. To see this, observe that when $\beta_2 < 1$, as long as we get self 2 to choose first-best savings (independent of whether we need to set $(1-\tau_2) > 0$ or < 0 to achieve this), we have $\hat{c}_2^* < c_2^*$. Hence self 1 will think that there is too little consumption in period two and hence will oversave. Moreover, if $\beta_2 = 1$, then future component of self 1's tax will be zero independent of the value of τ_2. So, again what matters for future component of self 1's tax is not future government policy but next period self's degree of self-control problem.

We have shown that the future component of the tax is positive under both full sophistication and full naivete. Since the future component under partial sophistication is a weighted average of the two, we have that for any π_1 future component is positive:

$$\left\{\pi_1\left[1+b_2'(b_1)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1^*)}\right]+(1-\pi_1)\frac{u'(\hat{c}_2)}{u'(c_2)}\right\}^{-1}=\left\{\pi_1\left(1-\tau_1^S\right)^{-1}+(1-\pi_1)\left(1-\tau_1^N\right)^{-1}\right\}^{-1}<1.$$

Equivalence under Logarithmic Utility

It is relatively easy to show that if the utility function is logarithmic, then:

$$b_2(b_1)=\left(R(1-\tau_1)b_1+T_1+\frac{T_2}{R(1-\tau_2)}\right)\frac{\beta_2\delta}{1+\beta_2\delta}.$$

Sophisticated future component:

$$(1-\tau_1^S)=\left[1+R(1-\tau_1^S)\frac{\beta_2\delta}{1+\beta_2\delta}\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1^S)}\right]^{-1}=\frac{1+\beta_2\delta}{1+\delta}$$

Naive future component:

$$(1-\tau_1^N)=\frac{\hat{c}_2^*}{c_2^*}=\frac{\left(R(1-\tau_1)b_1+T_1+\frac{T_2}{R(1-\tau_2)}\right)\frac{1}{1+\delta}}{\left(R(1-\tau_1)b_1+T_1+\frac{T_2}{R(1-\tau_2)}\right)\frac{1}{1+\beta_2\delta}}=\frac{1+\beta_2\delta}{1+\delta}.$$

We have just seen that when the utility is logarithmic, the future component of the optimal tax is the same for fully naive and fully sophisticated agents. Therefore, partial sophistication future component is independent of the degree of sophistication, π_1 :

$$\left\{\pi_1\left[\frac{1+\delta}{1+\beta_2\delta}\right]+(1-\pi_1)\frac{1+\delta}{1+\beta_2\delta}\right\}^{-1}=\frac{1+\beta_2\delta}{1+\delta}.$$

8 Appendix B - Proofs

8.1 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 first-best levels of capital stocks that start from k_0. We know that the first best 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 first-best:

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

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 first-best 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 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 behavior 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 intent $k_{t+s}^* = K(K(...(k_t^*)...))$, where the 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 first best. 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

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

$$\begin{aligned}
& && \text{s.t. for all } b_0 \\
b_1(b_0,k^1;\tau) &\in && \arg\max_{\hat{b}_1}\ u\left(R(k^1)\left(1-\tau_0\right)b_0+w(k^1)+T_0-\hat{b}_1\right) \\
& && +\delta\beta_1\left[\begin{array}{c}\pi_1\left\{\sum_{i=1}^{I-1}\delta^{i-1}u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(\hat{b}_1)..)+w(k^{i+1})+T_i-b_{i+1}(...(\hat{b}_1)..)\right)+\delta^{I-1}V\left(b_I(...(\hat{b}_1)..),k^{I+1};\tau\right)\right\} \\ +\left(1-\pi_1\right)W_1\left(\hat{b}_1,k^2;\tau\right)\end{array}\right] \\
& && \text{s.t. for all } b_1 \\
b_2(b_1,k^2;\tau) &\in && \arg\max_{\hat{b}_2}\ u\left(R(k^2)\left(1-\tau_1\right)b_1+w(k^2)+T_1-\hat{b}_2\right) \\
& && +\delta\beta_2\left[\begin{array}{c}\pi_2\left\{\sum_{i=2}^{I-1}\delta^{i-2}u\left(R(k^{i+1})\left(1-\tau_i\right)b_i(...(\hat{b}_2)..)+w(k^{i+1})+T_i-b_{i+1}(...(\hat{b}_2)..)\right)+\delta^{I-2}V\left(b_I(...(\hat{b}_2)..),k^{I+1};\tau\right)\right\} \\ +\left(1-\pi_2\right)W_2\left(\hat{b}_2,k^3;\tau\right)\end{array}\right] \\
& && \text{s.t. for all } b_2 \\
& && ... \\
& && \text{s.t. for all } b_{I-2} \\
b_{I-1}(b_{I-2},k^{I-1};\tau) &\in && \arg\max_{\hat{b}_{I-1}}\ u\left(R(k^{I-1})\left(1-\tau_{I-2}\right)b_{I-2}+w(k^{I-1})+T_{I-2}-\hat{b}_{I-1}\right)+\delta\beta_{I-1}\left(1-\pi_{I-1}\right)W_{I-1}\left(\hat{b}_{I-1},k^I;\tau\right) \\
& && +\delta\beta_{I-1}\left[\pi_{I-1}\left\{u\left(R(k^I)\left(1-\tau_{I-1}\right)\hat{b}_{I-1}+w(k^I)+T_{I-1}-b_I(...(\hat{b}_{I-1})..)\right)+\delta V\left(b_I(...(\hat{b}_{I-1})..),k^{I+1};\tau\right)\right\}\right] \\
& && \text{s.t. for all } b_{I-1} \\
b_I(b_{I-1},k^I;\tau) &\in && \arg\max_{\hat{b}_I}\ u\left(R\left(1-\tau_{I-1}\right)b_{I-1}+w_{I-1}+T_{I-1}-\hat{b}_I\right)+\delta\beta_I\left[\pi_I V\left(\hat{b}_I,k^{I+1};\tau\right)+\left(1-\pi_I\right)W_I\left(\hat{b}_I,k^{I+1};\tau\right)\right]
\end{aligned}$$

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

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

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):

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

where $G_i(k^{i+1})$ is the net present value of future lump-sum taxes 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 capital stock today is k^{i+1}, the level of assets is b_i 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, his 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 from

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

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:

$$\Gamma_i(b_i(b_{i-1}, k^i; \tau), k^{i+1}) = R(k^{i+1})(1 - \tau_i(k^{i+1}))(1 - M_i)\Gamma_{i-1}(b_{i-1}, k^i). \tag{8}$$

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

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

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:

$$\begin{aligned}
V(b, k; \tau) &= \max_{M_0} \; u(M_0 \Gamma_I(b)) + \delta \left[\sum_{i=1}^{I-1} \delta^i u(Q_{i-1}(k) M_i \Gamma_I(b)) + \delta^I V\left((1 - M_I) Q_{I-1}(k) \Gamma_I(b), k^{I+1}; \tau\right) \right] \qquad (9) \\
& s.t. \quad \text{for all } i \in \{1, ..., I-1\} \\
(M_i Q_{i-1}(k)\Gamma_I(b,k))^{-\sigma} \quad & \delta\beta_i \left[\begin{array}{c} \pi_i R(k^{i+1})(1 - \tau_i(k^{i+1})) \left\{ \begin{array}{c} \sum_{j=i+1}^{I} \delta^{j-(i+1)} \left(M_j Q_{j-1}(k) \Gamma_I(b,k)\right)^{-\sigma} M_j \frac{Q_{j-1}(k)}{Q_i(k)} \\ + \delta^{I-i} V'(b_I(..(b)..), k^{I+1}; \tau)(1 - M_I) \frac{Q_{I-1}(k)}{Q_i(k)} \end{array} \right\} \\ + (1 - \pi_i) W_i'\left(b_i(..(b)..), k^{i+1}; \tau\right) \end{array} \right] \\
(M_I Q_{I-1}(k)\Gamma_I(b,k))^{-\sigma} &= \delta\beta_I \left[\pi_I V'\left(b_I(..(b)..), k^{I+1}; \tau\right) + (1 - \pi_I) W_I'\left(b_I(..(b)..), k^{I+1}; \tau\right)\right].
\end{aligned}$$

Core proof of Proposition 1.

We will prove that facing the sequence of first-best capital stocks and the taxes specified in Proposition 1, people will choose first-best allocation, thereby verifying both (1) that the sequence of first-best capital stocks is actually part of equilibrium under Proposition 1 taxes, and (2) that under Proposition 1 taxes people choose first-best.

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):

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

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:

$$\begin{aligned} V'(b_I(..(b)..), k^{I+1}; \tau) &= DR(k^{I+1})(1-\tau_I(k^{I+1}))(\Gamma_I(b,k)Q_I(k))^{-1}, \qquad (10) \\ W_i'(b_i(..(b)..), k^{i+1}; \tau) &= D_i R(k^{i+1})(1-\tau_i(k^{i+1}))(\Gamma_I(b,k)Q_i(k))^{-1}, \end{aligned}$$

where we used the recursion (8).

Plugging these in the constraints described in problem (9), we get for all $i \in \{1, ..., I-1\}$:

$$\begin{aligned} (M_i Q_{i-1}(k))^{-1} &= \delta\beta_i R(k^{i+1})(1-\tau_i(k^{i+1}))\,(Q_i(k))^{-1} \left[\begin{array}{c} \pi_i \left\{\sum_{j=i+1}^{I} \delta^{j-(i+1)} + \delta^{I-i} D\right\} \\ + (1-\pi_i)\, D_i \end{array}\right] \\ & \textit{and} \\ (M_I Q_{I-1}(k))^{-1} &= \delta\beta_I R(1-\tau_I(k^{I+1}))\,(Q_I(k))^{-1}\left[\pi_I D + (1-\pi_I)\, D_I\right]. \end{aligned}$$

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

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

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 (9) 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 :

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

which implies

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

and hence

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

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

$$M_i=\frac{1-\delta}{1-\delta+\beta_i\delta},\ \text{for all } i\in\{1,..,I\},\tag{11}$$
$$M_0=1-\delta.$$

Now we turn to taxes that implement first-best allocation. The constraint that describes self-i's behavior for $i\in\{1,..,I-1\}$ becomes the following once we plug in the derivatives of the value functions from (10) :

$$\left(M_iQ_{i-1}(k)\Gamma_I\left(b,k\right)\right)^{-1}=\delta\beta_iR(k^{i+1})(1-\tau_i(k^{i+1}))\left(M_{i+1}Q_i(k)\Gamma_I\left(b,k\right)\right)^{-1}\left[\begin{array}{c}\pi_i\left\{\sum_{j=i+1}^{I}\delta^{j-(i+1)}+\delta^{I-i}D\right\}\\+\left(1-\pi_i\right)D_i\end{array}\right]M_{i+1}.\tag{12}$$

The comparison of (12) with the first-best condition (1) gives the first-best tax as:

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

For self-I, the constraint describing his behavior in problem (9) reads as follows:

$$\left(M_IQ_{I-1}(k)\Gamma_I\left(b,k\right)\right)^{-1}=\delta\beta_IR(k^{I+1})(1-\tau_I(k^{I+1}))\left(M_0Q_I(k)\Gamma_I\left(b,k\right)\right)^{-1}\left[\pi_ID+\left(1-\pi_I\right)D_I\right]M_0,$$

and the comparison of this with the first-best 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

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

with the corresponding first-best condition gives

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

8.2 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 first-best 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:

$$(1-\tau_i^*) = \frac{1}{\beta_i}\frac{1}{1+(1-M_{i+1}^*)\left(\frac{1}{\beta_{i+1}}-1\right)}. \tag{13}$$

When $R\delta = 1$, in the first best 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

$$M_i^* = \frac{M_{i+1}^*R(1-\tau_i^*)}{1+M_{i+1}^*R(1-\tau_i^*)}. \tag{14}$$

Plugging (13) in (14), we get a system of $(I+1)$ equations in $(I+1)$ unknows $(M_0^*, ..., M_I^*)$ that fully pin down agents policies when they face first-best 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 (11) satisfies the above system of equations, meaning it is an equilibrium. Plugging (11) in the formula for taxes, (13), we get that first-best taxes are the same as the logarithmic utility case.

9 Appendix C - Computational Procedure

9.1 Guess:

Guess

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

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.

9.2 Characterizing 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 $i \in \{0, 1, .., I-1\}$

$$
\begin{aligned}
D_i &= \left[\frac{[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1}{\sigma}} R(1-\tau_{i+1})}{1+[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1}{\sigma}} R(1-\tau_{i+1})}\right]^{1-\sigma}\left(1+\delta\frac{D_{i+1}}{[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1-\sigma}{\sigma}}}\right), \qquad (15)\\
D_I &= \left[\frac{[\delta R(1-\tau_0)D_0]^{-\frac{1}{\sigma}} R(1-\tau_0)}{1+[\delta R(1-\tau_0)D_0]^{-\frac{1}{\sigma}} R(1-\tau_0)}\right]^{1-\sigma}\left(1+\delta\frac{D_0}{[\delta R(1-\tau_0)D_0]^{-\frac{1-\sigma}{\sigma}}}\right).
\end{aligned}
$$

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

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

$$D = M_0^{-\sigma}. \qquad (16)$$

9.3 Characterizing optimal tax system, τ^*:

The incentive constraints for agents $i = 1, ..I$ together with parent's optimality condition with respect to bequest decision characterize 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 first-best euler equations, we immediately see that optimal taxes must satisfy:

$$\begin{aligned}
\text{For all } i &\in \{0,..,I-2\}, \qquad (17)\\
\left(1-\tau_{i+1}^{*}\right) &= \frac{1}{\beta_{i+1}}\left(\frac{\left[\pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{i-(i+2)}\left(M_j^{*}\frac{Q_{j-1}^{*}}{Q_{i+1}^{*}}\right)^{1-\sigma}+\delta^{I-(i+1)}D^{*}\left(\frac{Q_I^{*}}{Q_{i+1}^{*}}\right)^{1-\sigma}+\left(1-\pi_{i+1}\right)D_{i+1}^{*}\right\}\right]}{M_{i+2}^{*-\sigma}}\right)^{-1}\\
\left(1-\tau_I^{*}\right) &= \frac{1}{\beta_I}\left(\frac{\left[\pi_I D^{*}+\left(1-\pi_I\right)D_I^{*}\right]}{M_0^{*-\sigma}}\right)^{-1}\\
\left(1-\tau_0^{*}\right) &= \left(\frac{\left[\sum_{i=1}^{I}\delta^{i-1}\left(M_i\frac{Q_{i-1}^{*}}{Q_0^{*}}\right)^{1-\sigma}+\delta^{I}D^{*}\left(\frac{Q_I^{*}}{Q_0^{*}}\right)^{1-\sigma}\right]}{M_1^{*-\sigma}}\right)^{-1},
\end{aligned}$$

where D^* and D_i^* are first-best values computed according to (16) and (15) evaluated at first-best taxes.

9.4 Iteration

1. Before starting the iteration, compute first-best 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 first-best 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:

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

   (a) where

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

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

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

4. Compute D and $D_1, ..D_I$ according to (16) and (15) 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 (17).
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.

10 Appendix D - Partial Sophistication, CEIS coefficient, and Optimal Taxes

In this section, we formally answer the question: *Do taxes increase or decrease with increasing sophistication?* We illustrate this through the simple three period example we use in section 3. Since the degree of partial sophistication affects optimal taxes only through the future component, without loss of generality we assume $\beta_1 = 1$.

Self 1's problem:

$$\begin{aligned}
& \max_{b_1} u(k_0 - b_1) + \pi_1\delta\left[u(R(1-\tau_1)b_1 + T_1 - b_2(b_1)) + \delta u(R(1-\tau_2)b_2(b_1) + T_2)\right] \\
& +(1-\pi_1)\delta\left[u(R(1-\tau_1)b_1 + T_1 - \hat{b}_2(b_1)) + \delta u(R(1-\tau_2)\hat{b}_2(b_1) + T_2)\right] \\
& s.t. \\
b_2(b_1) &= \arg\max_{\tilde{b}_2} u(R(1-\tau_1)b_1 + T_1 - \tilde{b}_2) + \beta_2\delta u(R(1-\tau_2)\tilde{b}_2 + T_2) \\
\hat{b}_2(b_1) &= \arg\max_{\tilde{b}_2} u(R(1-\tau_1)b_1 + T_1 - \tilde{b}_2) + \delta u(R(1-\tau_2)\tilde{b}_2 + T_2)
\end{aligned}$$

Proposition 6 *If $\sigma > 1$ (< 1), then $\frac{\partial\tau_1^*}{\partial\pi_1} > 0 (< 0)$.*

Proof. The first order condition of self 1 reads:

$$\begin{aligned}
u'(c_1) &= \beta_1\delta\{\pi_1\left[u'(c_2)R(1-\tau_1) + b_2'(b_1)\{-u'(c_2) + \delta R(1-\tau_2)u'(c_3)\}\right] \\
&\quad +(1-\pi_1)\left[u'(\hat{c}_2)R(1-\tau_1) + \hat{b}_2'(b_1)\{-u'(\hat{c}_2) + \delta R(1-\tau_2)u'(\hat{c}_3)\}\right]\}
\end{aligned}$$

When utility is CEIS, a self 2 that has self-control problem level $\bar{\beta}_2$ consumes according to:

$$c_2(b_1) = M_2(\bar{\beta}_2)\left(R(1-\tau_1)b_1 + T_1 + \frac{T_2}{R(1-\tau_2)}\right),$$

where

$$M_2(\bar{\beta}_2) = \frac{R(1-\tau_2)}{R(1-\tau_2) + \left(R(1-\tau_2)\bar{\beta}_2\delta\right)^{\frac{1}{\sigma}}}.$$

With probability π_1 self 1 correctly believes that $\bar{\beta}_2 = \beta_2$ and with probability $1-\pi_1$, he believes that self 2 has no self-control problems and hence $\bar{\beta}_2 = 1$.

Now it follows from the budget that:

$$b_2(b_1) = (R(1-\tau_1)b_1 + T_1)(1 - M_2(\bar{\beta}_2)) - M_2(\bar{\beta}_2)\frac{T_2}{R(1-\tau_2)}$$

which implies

$$b_2'(b_1) = R(1-\tau_1)(1 - M_2(\bar{\beta}_2))$$

Plugging this in the first order condition, we get:

$$u'(c_1) = \beta_1 \delta R(1-\tau_1)u'(c_2)\left\{\begin{array}{c} \pi_1 \frac{[M_2(\beta_2)u'(c_2)+(1-M_2(\beta_2))\delta R(1-\tau_2)u'(c_3)]}{u'(c_2)} \\ +(1-\pi_1)\frac{[M_2(1)u'(\hat{c}_2)+(1-M_2(1))\delta R(1-\tau_2)u'(\hat{c}_3)]}{u'(c_2)} \end{array}\right\}.$$

So, the first best tax is:

$$1 - \tau_1^* = \frac{1}{\beta_1}\left\{\pi_1 \frac{\kappa(\beta_2)}{(c_2^*)^{-\sigma}} + (1-\pi_1)\frac{\kappa(1)}{(c_2^*)^{-\sigma}}\right\}^{-1}$$

where

$$\kappa(\bar{\beta}_2) = \left[M_2(\bar{\beta}_2)(c_2)^{-\sigma} + (1 - M_2(\bar{\beta}_2))\delta R(1-\tau_2)(c_3)^{-\sigma}\right] > 0.$$

Remember we want to compute:

$$\begin{aligned} sign\left(\frac{\partial(\tau_1^*)}{\partial \pi_1}\right) &= sign\left(\frac{\partial\left[\pi_1 + (1-\pi_1)\frac{\kappa(1)}{\kappa(\beta_2)}\right]}{\partial \pi_1}\right) \\ &= sign\left(1 - \frac{\kappa(1)}{\kappa(\beta_2)}\right) \end{aligned}$$

since first-best allocation is independent of π_1.

Now we show that $\frac{\kappa(1)}{\kappa(\beta_2)}$ is greater (smaller) than 1 when $\sigma < (>) 1$. One can show that after plugging in $M_2(\bar{\beta}_2)$, getting rid of period 1 wealth, and regrouping:

$$\frac{\kappa(\bar{\beta}_2)}{\kappa(\beta_2)} = \frac{\left[\left(\frac{R(1-\tau_2)}{R(1-\tau_2)+\left(R(1-\tau_2)\bar{\beta}_2\delta\right)^{\frac{1}{\sigma}}}\right)^{1-\sigma}\left\{1 + \delta\left(R(1-\tau_2)\bar{\beta}_2\delta\right)^{\frac{1}{\sigma}-1}\right\}\right]}{\left[\left(\frac{R(1-\tau_2)}{R(1-\tau_2)+(R(1-\tau_2)\beta_2\delta)^{\frac{1}{\sigma}}}\right)^{1-\sigma}\left\{1 + \delta\left(R(1-\tau_2)\beta_2\delta\right)^{\frac{1}{\sigma}-1}\right\}\right]}.$$

Then,

$$\frac{\partial \frac{\kappa(\bar{\beta}_2)}{\kappa(\beta_2)}}{\partial \bar{\beta}_2} = \kappa(\beta_2)^{-1}\left[\begin{array}{c} \frac{(1-\sigma)}{\sigma}\left(\frac{R(1-\tau_2)}{R(1-\tau_2)+\left(R(1-\tau_2)\bar{\beta}_2\delta\right)^{\frac{1}{\sigma}}}\right)^{1-\sigma} R(1-\tau_2)\delta\left(R(1-\tau_2)\bar{\beta}_2\delta\right)^{\frac{1}{\sigma}-1} \\ \left(\frac{1}{\bar{\beta}_2} - 1\right)\left\{\frac{1}{R(1-\tau_2)\left(R(1-\tau_2)+\left(R(1-\tau_2)\bar{\beta}_2\delta\right)^{\frac{1}{\sigma}}\right)}\right\} \end{array}\right],$$

which implies

$$sign\frac{\partial \frac{\kappa(\bar{\beta}_2)}{\kappa(\beta_2)}}{\partial \bar{\beta}_2} = sign\frac{(1-\sigma)}{\sigma}.$$

This means that, for any π_1, if $\sigma > 1 (< 1)$, then for $\bar{\beta}_2 \in (\beta_2, 1)$, $\frac{\kappa(\bar{\beta}_2)}{\kappa(\beta_2)}$ is increasing (decreasing) with $\bar{\beta}_2$, which further implies $\frac{\kappa(1)}{\kappa(\beta_2)} < (>) 1$ since $\frac{\kappa(\beta_2)}{\kappa(\beta_2)} = 1$. Thus, we get the following result:

$$sign\left(\frac{\partial(\tau_1^*)}{\partial \pi_1}\right) > (<) 0 \text{ if } \sigma > (<) 1.$$

11 Appendix E - Introducing an Illiquid Asset

Consider again the three period example of Section 3, 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 first-best taxes in the environment without illiquid asset, τ_1^*, τ_2^*, we implement first-best allocation in the market with the illiquid asset. Let us compute these taxes for future use. Since

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

first-best 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 $\frac{\partial y_1(b_1,d_1)}{\partial b_1} = R(1-\tau_1)$. 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 first-best net present value of wealth.

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, first-best still requires

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

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:

$$\begin{aligned} R^d(1-\tau^{d*}) &= R(1-\tau_2^*)\frac{u'(c_1^*)}{\delta R\left(\pi_1\left[u'(c_2^*)c_2'(y_1^*) + \delta u'(c_3^*)c_3'(y_1^*)\right] + (1-\pi_1)\left[u'(\hat{c}_2^*)\hat{c}_2'(y_1^*) + \delta u'(\hat{c}_3^*)\hat{c}_3'(y_1^*)\right]\right)} \\ &= R(1-\tau_2^*)R(1-\tau_1^*). \end{aligned} \quad (18)$$

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 (18). In either case, the taxes on period one and period two liquid assets are exactly equal to the first-best taxes in the environment without illiquid assets.