

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

Nicola Pavoni[†] and Hakki Yazici[‡]

November 2013

Abstract

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

JEL classification: E21, E62, D03.

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

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

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

[‡]Sabanci University.

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

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 on self-control problems is that the degree of self-control problem is constant over time. This contrasts with the significant body of empirical research indicating that, like many other personality traits, people's ability to self-control changes as they age. A 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

¹See 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.

²Three 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).

³Actually, 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."

with age.⁴ Indeed, Caspi, Roberts, Robins, and Trzesniewski (2003), in their survey article on personality development in adulthood, conclude that: "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 In this paper, we extend the traditional models of self-control to allow for varying degrees of self-control problem over the life cycle, and study optimal capital income taxation.

In our model, agents make consumption and savings decisions facing self-control problems at all ages. In the last period of their lives, people make consumption and bequest decisions knowing that they are going to be replaced by their offspring next period. We model preferences that exhibit self-control problems through the quasi-hyperbolic discounting framework of Laibson (1997), 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 future self-control problems.

In this environment, we define *efficient (or commitment)* allocation as the allocation that would arise in the absence of self-control problems. It is given by the solution to a fictitious social planner's consumption-saving problem where the *planner discounts exponentially* future utilities. In our environment, this preference corresponds to the preference of an initial generation parent. The main exercise in this paper is to examine the optimal tax policy that implements the efficient allocation. In this sense, this paper is a normative exploration of optimal paternalistic tax policy regarding life-cycle saving behavior. It is well-known that in models

⁴For 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.

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

of quasi-hyperbolic discounting there is multiplicity of equilibria.⁶ We restrict attention to the (unique) linear Markov equilibrium of our economy.

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

Our closed forms are obtained assuming agents have CEIS preferences. 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 study the quantitative effects of age-dependent self-control in a calibrated version of our model, and we find that the optimal subsidies in our model with decreasing self-control problems are *much larger than those implied by a model with constant self-control*. We also compare our optimal subsidies with those implied by the 401(k) plan. If we exclude the very last periods before retirement - where the subsidy rate in the 401(k) essentially mimics the employer matching

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

rate - the subsidy levels in the two cases are of comparable magnitudes. A marked difference emerges however: the 401(k) plan implies an increasing pattern of subsidies while the optimal subsidies decrease over the life cycle.

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. The main conclusion from our numerical experiments is as follows: as long as the level of sophistication is not changing abruptly from one period to another, the pattern of optimal capital subsidies over the life cycle is surprisingly robust to the degree of sophistication. This result holds approximately for a large range of CEIS coefficients.

Related Literature. Our paper is related to a number of recent papers that have explored the implications of self-control problems for optimal paternalistic taxation. O'Donoghue and Rabin (2003b) and O'Donoghue and Rabin (2006) analyze models of paternalistic taxation of unhealthy goods. More closely related is Krusell, Kuruscu, and Smith (2010), which analyzes optimal taxation of savings in an economy where agents live finitely many periods and have temptation and self-control problems à la Gul and Pesendorfer (2001).⁸ First, they prove that the optimal policy prescriptions of the quasi-hyperbolic model and the temptation model are identical when the utility function is logarithmic or when it is CEIS and the temptation parameter goes to infinity. Second, they show that savings should be subsidized and that this *subsidy should be increasing with time* due to finite life time effect.⁹ Our work differs from this paper along several dimensions. First and foremost, we allow for changing level of self-control problems over the life cycle whereas - like all papers prior to ours - Krusell, Kuruscu, and Smith (2010)

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

⁸Krusell, Kuruscu, and Smith (2002) also analyze optimal taxation of savings under self-control problems but their main focus is on an environment where the government as well as the people face issues of time-inconsistency.

⁹It is indeed straightforward to show that, in the infinite horizon version of their model, the subsidies would be constant.

assume the level of self-control problem to be constant over time. The implications of modeling age-dependent self-control problems turns out to be significant. First, by assuming empirically plausible patterns of self-control problems over the life cycle, we show, analytically, that capital subsidies should actually be decreasing with age. Our benchmark model assumes perfectly altruistic parents, making it equivalent to a standard infinite horizon framework. In the quantitative section, we allow parents to be imperfectly altruistic and generalize our optimal tax formulas to take into account the finite life time effects of Krusell, Kuruscu, and Smith (2010). We find that, in our model, the effect of age-dependent self-control dominates the finite time effect induced by imperfect altruism: optimal capital income subsidies decrease over the life cycle even when parents do not care at all about their offsprings (i.e., when finite life time effect is the strongest). Second, we show quantitatively that age-dependence of self-control problems imply much higher levels of optimal subsidies relative to the constant self-control model. Finally, 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 an assumption maintained in Krusell, Kuruscu, and Smith (2010). This allows us to study the effects of sophistication on capital subsidies.

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

As discussed above, an immediate implication of age-dependent self-control problems is that capital taxes should be age-dependent. The age-dependence result is also a feature of two sets of earlier contributions that 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., Farhi and Werning (2013) and Golosov, Troshkin, and Tsyvinski (2011)). The forces generating age-dependence in the current paper, however, are completely different from the forces in these papers.¹⁰ Therefore, our paper complements this literature by providing a new mechanism through which capital taxes should depend on age. As such, in this paper, the life-cycle pattern of optimal capital taxes depends on features of the environment that are neglected by these papers.

2 Model

The economy is populated by a continuum of a unit measure of dynasties who live for a countable infinity of periods, $t = 1, 2, \dots$, 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.¹¹ We use quasi-hyperbolic discounting formalized by Laibson (1997) to

¹⁰The optimality of age-dependence in Erosa and Gervais (2002) is a direct implication of time-dependent consumption and labor plans present in Ramsey tax models off the steady-state. In the New Dynamic Public Finance models, capital is taxed in order to deter people from joint deviation of saving and shirking. Since people at different ages (and contingencies) have different levels of accumulated wealth and future prospects, they have different tendencies to save, and hence, the corrective taxes depend on age.

¹¹ 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 'perfectly' altruistic parents) to abstract away from the effects of finite life time on taxes (see Krusell, Kuruscu, and Smith (2010) for finite life time effects). The case with imperfect altruism is briefly analyzed

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:

$$u(c_0) + \delta u(c_1) + \delta^2 u(c_2) + \dots + \delta^I u(c_I) + \delta^{I+1} u(c'_0) + \dots$$

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 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:

$$\begin{aligned} &u(c_1) + \beta_1 \delta \left[u(c_2) + \delta u(c_3) + \dots + \delta^{I-1} u(c'_0) + \dots \right], \\ &u(c_2) + \beta_2 \delta \left[u(c_3) + \dots + \delta^{I-2} u(c'_0) + \dots \right], \\ &\dots, \\ &u(c_I) + \beta_I \delta \left[u(c'_0) + \dots \right]. \end{aligned}$$

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.

in Section 4.1.

We allow for partial sophistication which essentially amounts to allowing for different degrees of self awareness about the existence of self-control problems.¹² 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 $\sigma > 0$:

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

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:

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

Labor is inelastically supplied, so at all dates $l = 1$. Define

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

where d refers to the fraction of capital that is forgone due to depreciation. There is a credit market in which agents can trade one period risk-free bonds and capital as perfectly substitutable assets. Since at any given date there is not cross-sectional heterogeneity, all agents have the same level of asset holdings. 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 The Efficient Allocation

The *efficient* or – as we use interchangeably throughout the paper – the *commitment* allocation is the allocation that would arise in the absence of self-control problems. It is given by the so-

¹²We are not the first ones to model partial sophistication, O'Donoghue and Rabin (1999) are. However, the way we introduce partial sophistication is different from theirs, and more in line with Eliaz and Spiegel (2006) and Asheim (2007). We justify our way of modeling partial sophistication on the grounds of tractability. The added bonus of our model of partial sophistication is that the structure is consistent with a learning approach to sophistication (e.g., Ali (2011)).

lution to a fictitious social planner’s consumption-saving problem where the *planner discounts exponentially* with discount factor δ . In our environment, this preference corresponds to the preference of an initial generation parent. By taking a long-term perspective and evaluating welfare according to the initial generation parent’s preference, we are following much of the literature.¹³ The following Euler Equations characterize the efficient allocation, which we denote with an asterisk throughout the paper:¹⁴

$$u'(c_i^*) = \delta f'(k_i^*) u'(c_{i+1}^*), \text{ for } i = 0, 1, 2, \dots, I - 1, \quad (1)$$

and

$$u'(c_I^*) = \delta f'(k_I^*) u'(c_0^*),$$

...

2.2 Implementing the Efficient Allocation

Since people in this economy face self-control problems, laissez-faire market equilibrium cannot attain the commitment allocation. Our main interest in this paper is to find and characterize a tax system that implements the commitment allocation in the market environment. We call such a tax system *optimal*. We proceed by defining a market equilibrium with taxes. It is important to note that from the outset we restrict the set of taxes that are available to the government to linear taxes on savings coupled with lump-sum rebates (throughout the paper we call this the set of linear taxes). In general, it is not obvious that there is a linear tax system that implements the efficient allocation. However, since we focus our attention to linear equilibria, a linear tax system that implements the efficient allocation exists. We will verify this claim in Section 3.

¹³See DellaVigna and Malmendier (2004), Gruber and Koszegi (2004) and O’Donoghue and Rabin (2006), for example.

¹⁴We do not state the transversality condition but the commitment allocation will converge to a steady state with positive capital as long as $k_0 > 0$.

2.3 Markov Equilibrium with Taxes

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

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

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

Let τ_i be the savings (capital) tax agent $i = 0, 1, \dots, I$ 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 $i = 0, 1, \dots, I$, 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 $b_i(b_{i-1}(\dots b_{n+1}(b_n; \tau)\dots; \tau); \tau)$, which will be referred to as $b_i(\dots(b_n)\dots)$ 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.

In order to define the equilibrium for this economy, we first define the parent's problem. Let $V(b; \tau)$ be 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, \dots, 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.

Problem of the parent:

$$\begin{aligned}
V(b; \tau) &= \max_{b_0} u(R(1 - \tau_I)b + w + T_I - b_0) \\
&\quad + \delta \left\{ \sum_{i=0}^{I-1} \delta^i u(R(1 - \tau_i)b_i(\dots(b_0)\dots) + w + T_i - b_{i+1}(\dots(b_0)\dots)) + \delta^I V(b_I(\dots(b_0)\dots); \tau) \right\} \\
&\text{s.t. for all } b_0 \\
&\quad b_1(b_0; \tau) = \\
&\arg \max_{\hat{b}_1} u\left(R(1 - \tau_0)b_0 + w + T_0 - \hat{b}_1\right) + \delta\beta_1 \left[\pi_1 u\left(R(1 - \tau_1)\hat{b}_1 + w + T_1 - b_2(\hat{b}_1)\right) + (1 - \pi_1) W_1(\hat{b}_1; \tau) \right] \\
&\quad + \delta\beta_1 \pi_1 \left\{ \sum_{i=2}^{I-1} \delta^{i-1} u\left(R(1 - \tau_i)b_i(\dots(\hat{b}_1)\dots) + w + T_i - b_{i+1}(\dots(\hat{b}_1)\dots)\right) + \delta^{I-1} V\left(b_I(\dots(\hat{b}_1)\dots); \tau\right) \right\} \\
&\text{s.t. for all } b_1 \\
&\quad b_2(b_1; \tau) = \\
&\arg \max_{\hat{b}_2} u\left(R(1 - \tau_1)b_1 + w + T_1 - \hat{b}_2\right) + \delta\beta_2 \left[\pi_2 u\left(R(1 - \tau_2)\hat{b}_2 + w + T_2 - b_3(\hat{b}_2)\right) + (1 - \pi_2) W_2(\hat{b}_2; \tau) \right] \\
&\quad + \delta\beta_2 \pi_2 \left\{ \sum_{i=3}^{I-1} \delta^{i-2} u\left(R(1 - \tau_i)b_i(\dots(\hat{b}_2)\dots) + w + T_i - b_{i+1}(\dots(\hat{b}_2)\dots)\right) + \delta^{I-2} V\left(b_I(\dots(\hat{b}_2)\dots); \tau\right) \right\} \\
&\quad \dots \\
&\quad b_{I-1}(b_{I-2}; \tau) = \\
&\arg \max_{\hat{b}_{I-1}} u\left(R(1 - \tau_{I-2})b_{I-2} + w + T_{I-2} - \hat{b}_{I-1}\right) + \delta\beta_{I-1} \pi_{I-1} u\left(R(1 - \tau_{I-1})\hat{b}_{I-1} + w + T_{I-1} - b_I(\hat{b}_{I-1})\right) \\
&\quad + \delta\beta_{I-1} \left[(1 - \pi_{I-1}) W_{I-1}(\hat{b}_{I-1}; \tau) + \delta\pi_{I-1} V\left(b_I(\hat{b}_{I-1}); \tau\right) \right] \tag{3}
\end{aligned}$$

$$\begin{aligned}
&\text{s.t. for all } b_{I-1} \\
&\quad b_I(b_{I-1}; \tau) = \arg \max_{\hat{b}_I} u\left(R(1 - \tau_{I-1})b_{I-1} + w + T_{I-1} - \hat{b}_I\right) + \delta\beta_I \left[\pi_I V\left(\hat{b}_I; \tau\right) + (1 - \pi_I) W_I\left(\hat{b}_I; \tau\right) \right] \tag{4}
\end{aligned}$$

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

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

with

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

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 above; (iii) the government budget is satisfied period-by-period and markets clear: $T_i = R\tau_i b_i(k; \tau)$ and $b_i(k; \tau) = k$ for all i .

To understand the nested nature of policies and the way we model partial sophistication better, let us analyze the definition of policies in (3) and (4). First, constraint (4) 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. $(\beta_1, \dots, \beta_I) \neq (1, \dots, 1)$, 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. $(\beta_1, \dots, \beta_I) = (1, \dots, 1)$. 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.¹⁵ Second, consider constraint (3) 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 (4), 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, (4), enters the parent's problem in at least two ways: first, in the definition of self I 's policy function and then as a constraint in the

¹⁵Of 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.

definition of self $I - 1$'s policy function. These two different constraints are represented by a single constraint, (4), because the parent and self $I - 1$'s sophisticated belief agree about how self I will behave.¹⁶ Similarly, the constraint describing self $I - 1$'s policy is also a constraint in the constraint that describes self $I - 2$'s policy, and self $I - 2$'s policy is also a constraint of self $I - 3$'s, and so on. Thus, actually the constraint that describes the policy of self i enters parent's problem in i different places but since these are all identical constraints, we represent them with just one constraint that describes self i 's policy.

We restrict attention to linear equilibria, meaning equilibria 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

$$\Gamma_i(b) = R(1 - \tau_i)b + w + T_i + G_i,$$

$$G_i = \frac{T_{i+1} + w}{R(1 - \tau_{i+1})} + \frac{T_{i+2} + w}{R^2(1 - \tau_{i+1})(1 - \tau_{i+2})} + \dots + \frac{T_I + w}{R^{I-i} \prod_{j=i+1}^I (1 - \tau_j)} + \frac{T_0 + w}{R^{I-i+1}(1 - \tau_0) \prod_{j=i+1}^I (1 - \tau_j)} + \dots,$$

where G_i is the total net present value of future lump-sum taxes plus wages, $\Gamma_i(b)$ 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.

¹⁶Sophisticated belief of self i about how self $j, j > i$, agrees with parent's belief thanks to our assumption that the same 'beliefs' $(\pi_i)_i$ are shared by all agents.

3 Optimal Taxes

In this section we analyze optimal capital taxes in the model introduced in Section 2. Proposition 1 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, $\pi = (\pi_1, \pi_2, \dots, \pi_I)$, we have:*

$$1 - \tau_0^* = 1 - \delta + \beta_1 \delta,$$

$$1 - \tau_i^* = \frac{1}{\beta_i} (1 - \delta + \beta_{i+1} \delta), \text{ for } i \in \{1, \dots, I - 1\}$$

$$1 - \tau_I^* = \frac{1}{\beta_I}.$$

Proof. In Appendix A. ■

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.

It might be important to stress that - as shown in the Appendix - Proposition 1 holds regardless of whether the economy is in a steady state or in a transition. In particular, since agents do not face binding liquidity constraints, *the expressions for taxes hold for any life-cycle path of wages*: they only depend on the path of self-control parameter β_i . In Proposition 2, 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 , i.e., optimal capital taxes take exactly the same form as those of Proposition 1 for all i .*

Proof. In Appendix A. ■

The tax formula for $1 - \tau_i^*$, $0 < i < I$, consists of two main components. The first part, $\frac{1}{\beta_i}$, is easier to understand. Because of his *current* self-control problem, self i discounts tomorrow by an extra β_i and hence wants to undersave relative to the efficient allocation. By multiplying the after tax return with $\frac{1}{\beta_i}$, 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 is always greater than one, i.e. it always calls for a subsidy. This is not the end of the story, however. Self i '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 efficient saving level in order to compensate for his future self's suboptimal actions (due to future self-control problems) and/or in response to future policies. The component $(1 - \delta + \beta_{i+1}\delta)$ of the tax formula 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. Here, the level of sophistication might also matter. Interestingly, it can be shown that the future component is always less than one, i.e. *it calls for a tax, independent of the level of sophistication*.¹⁷

To gain intuition on why future component always calls for a tax, consider the future component in the case where agents are fully sophisticated, i.e. set $\pi_i = 1$ for all i . The future component is less than one because of a combination of two factors. First, it is easy to see that the policy of each agent in this economy $b_i(\cdot)$ is monotone in the level of assets. Second, from self i 's perspective self $i + 1$ undersaves in period $i + 1$. This implies a violation of an Envelope condition that holds when agents are time-consistent. According to agent i , each unit saved by self $i + 1$ has a cost $u'(c_{i+1}^*)$ that is lower than the self i 's perceived return. This is so since self $i + 1$'s discount rate between period $i + 1$ and $i + 2$ is higher than that of self i . Self $i + 1$'s undersaving hence appears as an extra return to saving for self i . As a consequence, whenever the policy function of agent $i + 1$ is monotone in the level of assets, self i is induced to save more than the efficient level. The future component of the tax formula calls for a tax in order to correct this oversaving behavior. The reader might still feel puzzled by our argument: after

¹⁷For details, the reader can refer to Section 3 of an earlier version of the paper: Pavoni and Yazici (2012).

all, when facing the optimal taxes, self $i + 1$ saves the efficient amount. Note however, that from self 1's perspective, self $i + 1$ is still undersaving (at the new price that is inflated by the subsidy).

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

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

3.1 Lessons for Capital Taxation

Propositions 1 and 2 imply 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. In particular, depending on how β_i changes with i :*

- (i) *Optimal capital taxes might be positive or negative at different ages.*
- (ii) *Optimal capital taxes might be increasing or decreasing with age at different ages.*

Proof. (i) For an example of $\tau_i > 0$, set $\beta_{i+1} \approx 0$ and $\beta_i > 1 - \delta$.

For an example of subsidy, set $\beta_i = \beta_{i+1} = \beta < 1$. See also Figure 1.

(ii) See the green line with crosses in Figure 2 below for an example. ■

Corollary 3 shows 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. It also shows that optimal capital taxes might be

positive or negative and that they can be increasing or decreasing with age, all depending on how the severity of the self-control problems evolve over the life cycle. These ambiguity results about the qualitative properties of optimal taxes constitute an important message: 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, for instance, to the presumption in the literature that self-control problems always imply subsidies.¹⁸ 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 the commitment allocation calls for subsidies.

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 Figure 1, we see that constant β_i , which is depicted by a dashed line, implies constant subsidies as found by previous literature. This figure also shows that optimal capital taxes can be positive if β_i decreases with age as depicted by the red crosses. In Figure 2, we see different self-control patterns that are all increasing with age. In all these cases, as the theory shows, capital should be subsidized. Whether subsidies should increase or decrease with age depends on the curvature of β_i .

Corollaries 4 and 5 below characterize quite sharply the sign and monotonicity properties of optimal capital taxes over the life cycle for a class of self-control patterns over the life cycle that is suggested by personality psychologists and the literature on intertemporal discounting.

The pattern is that the degree of self-control problem decreases with age and this decline slows down with age. In the notation of this paper, this means β_i is increasing and concave in i . 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

¹⁸O'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.

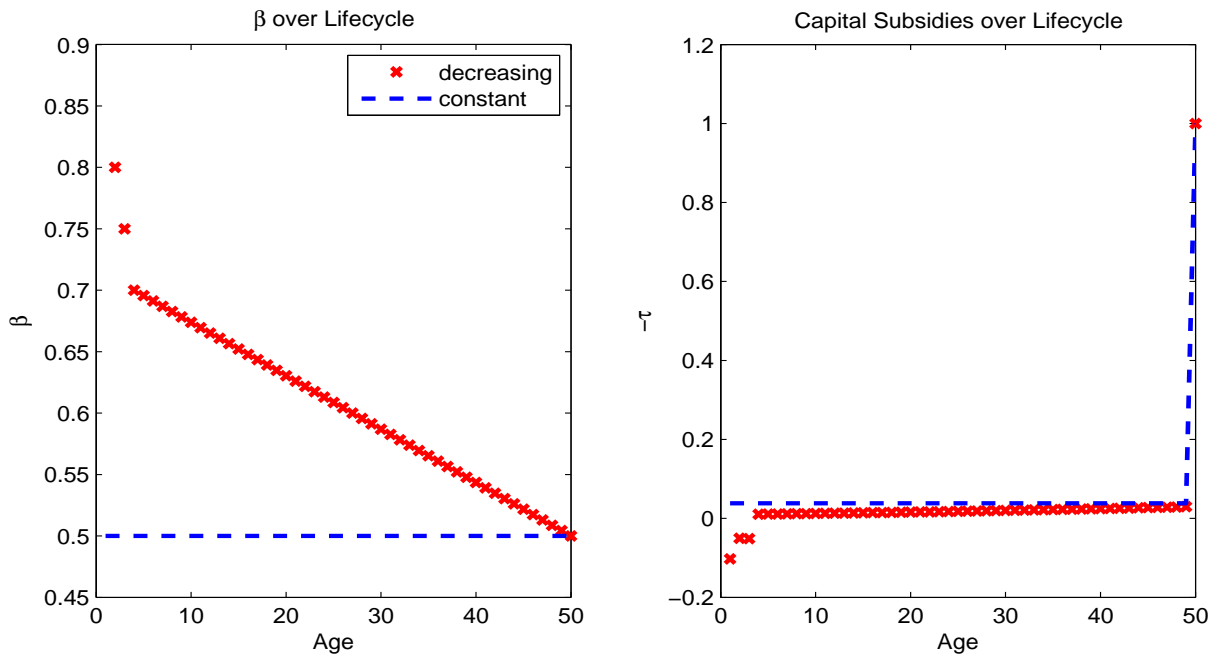


Figure 1: *Optimal capital subsidies for decreasing and constant patterns of $i \rightarrow \beta_i$ over the life cycle.*

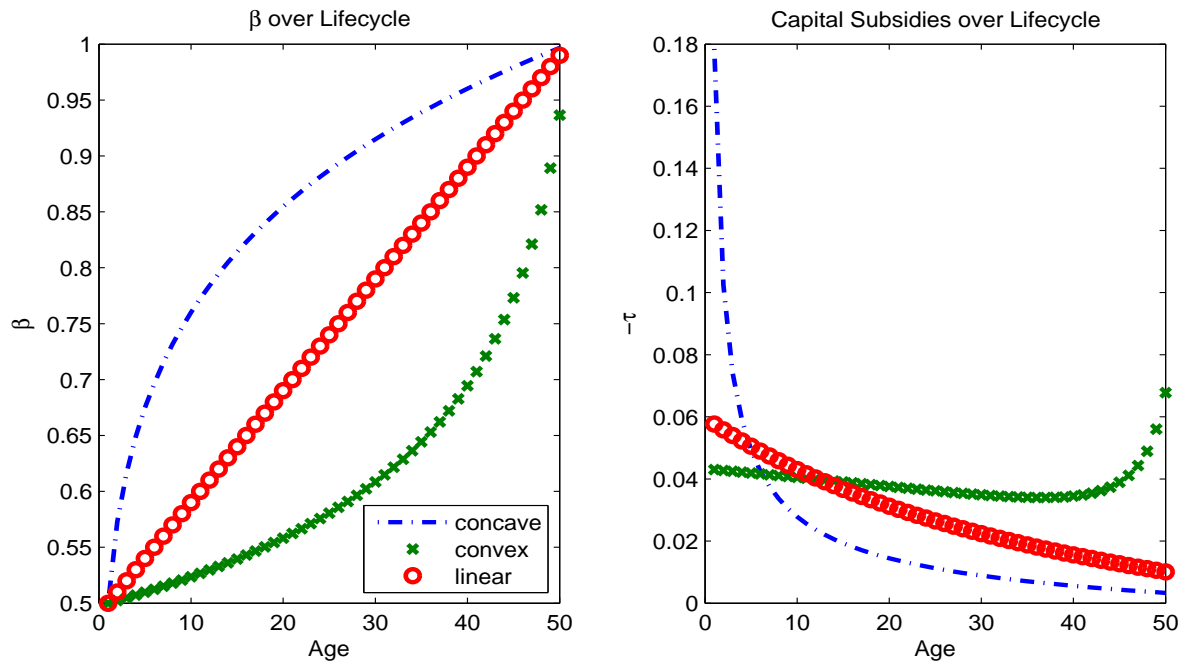


Figure 2: *Optimal capital subsidies for concave, linear, and convex increasing patterns of $i \rightarrow \beta_i$ over the life cycle.*

predicting a life-cycle developmental trend toward increased self-control.¹⁹ Second, personality 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. For example, in the work by John, Gosling, Potter, and Srivastava (2003), conscientiousness is estimated as a quadratic function of age and they find 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

Corollary 4 shows that if the severity of self-control problems decline with age, then capital should be subsidized at all ages.

Corollary 4 (*Optimality of Capital Subsidies*)

Under the assumptions of Propositions 1 or 2, if $\beta_{i+1} \geq \beta_i$, for all i , optimal capital tax is negative for all ages:

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

¹⁹See Green, Fry, and Myerson (1994), Green, Myerson, and Ostadzewski (1999), Read and Read (2004), and Ameriks, Caplin, Leahy, and Tyler (2007).

²⁰Ameriks, 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.

²¹It is possible to compute one-year short-term discount rates (our β 's) using Green, Myerson, and Ostadzewski (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, they have only three age groups.

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

Corollary 5 (Decreasing Capital Subsidies)

Under the assumptions of Propositions 1 or 2, if $0 \leq \beta_{i+1} - \beta_i \leq \beta_i - \beta_{i-1}$ for all i (concavity), then optimal capital subsidies decrease with age.

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

The result of Corollary 5 is contrary to Krusell, Kuruscu, and Smith (2010) who conclude that in any finite economy with constant self-control, capital subsidies should be increasing with age. The optimality of increasing subsidies in their case is purely due to the finite life time people face, and this element is missing from our analysis due to our assumption of perfect altruism. In Section 4.1, we show that the finite life time effect is quantitatively small for the relevant parameter space, implying that the optimality of decreasing subsidies with age is generally optimal.

3.2 Quantitative 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 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 with age, $\{\beta_i\}_{i=1}^{51}$.²² We set the

²²Observe 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

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 vector $\{\beta_i\}_{i=1}^{51}$ is the crucial 'parameter'. Figure 1 and Figure 2 show that taxes are in general very sensitive to the vector $\{\beta_i\}_{i=1}^{51}$. This vector is calibrated as follows. We assume the relationship $i \rightarrow \beta_i$ takes the following functional form:

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

This form allows for both concave and convex patterns. A priori, we do not want to restrict the sign of the curvature; therefore, we let the parameters to be dictated by the calibration targets. Note that whenever the signs of d and b are the same, $i \rightarrow \beta_i$ is an increasing function of i . We have chosen this functional form over some other - perhaps simpler - functions because it satisfies the key condition of our model, namely $\beta_i \leq 1$ for all i . In Figure 6, we provide some sensitivity analysis regarding the functional forms.

We use the level of self-control problems at the beginning and end of the life cycle, β_1 and β_{51} , together with the average level of self-control problems in the economy, call it β_{avg} , to pin down the parameters of the function. In our benchmark simulation, we assume $\beta_1 = 0.5$. Green, Myerson, and Ostaszewski (1999) estimates a hyperbolic discount function for their young adult group which has mean age of 20 years. We approximate a quasi-hyperbolic discount function using their estimate. Even though this approximation procedure does not allow us to pin down a single number for one-year short term discount rate for this group, β_1 , it does allow us to compute upper and lower bounds on the level of β_1 . Our benchmark choice of $\beta_1 = 0.5$ is in the middle of this range, $[0.25, 0.7]$.²³ We then perform sensitivity analysis for $\beta_1 = 0.35$ and $\beta_1 = 0.65$. We assume that self-control problem vanishes towards the end of

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.

²³In Appendix C, we explain in detail how we approximate the bounds on β_1 using Green, Myerson, and Ostaszewski (1999).

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. Thus, β_{51} is set to 1. We take β_{avg} to be 0.818 in our benchmark analysis following the estimate of Laibson, Repetto, and Tobacman (2007) for $\sigma = 1$ for a constant self-control model with mean age 40. We perform sensitivity analysis by also targeting β_{avg} to 0.703 and 0.898. The first value corresponds to Laibson, Repetto, and Tobacman (2007)'s estimates for $\sigma = 2$; our second target for β_{avg} is the value they find when they estimate σ and β jointly.²⁴

By construction, $\beta_{51} = 1$ implies $a = 1 + d$. Then, for a given level of d , we compute b as a function of d using equation (5) for $i = 1$. Finally, we pick d to match β_{avg} . The calibration for benchmark and sensitivity exercises are summarized in Table 1. The benchmark calibration implies a concave pattern for $i \rightarrow \beta_i$. In the robustness check exercises, we also obtain three calibrations with negative coefficients, implying a convex pattern of self-control.

Now we summarize our results. The solid blue line on the left-hand panel of Figure 3 represents our benchmark calibration of the evolution of self-control problem over the life cycle; the right-hand panel displays the corresponding optimal subsidies. In our benchmark exercise, optimal taxes are negative, so they are indeed subsidies. The subsidies start at 8% and are decreasing with age to less than 1% toward the end of the life cycle. The optimality of declining subsidies is expected given Corollary 4 and Corollary 5 and the concavely increasing pattern of β_i with i .

The dotted lines in Figure 3 display the results of the sensitivity analysis we performed regarding one of our calibration targets, β_1 . We recalibrate the $i \rightarrow \beta_i$ function given by equation (5) targeting first $\beta_1 = 0.35$ and then $\beta_1 = 0.65$. The calibrated parameters are reported in second and third rows of Table 1, respectively. We observe that when we calibrate the evolution

²⁴Recall that our tax formulas are valid for any level of σ if the steady-state condition specified in Proposition 2 holds.

Table 1: Calibration of self-control pattern

Targets			Parameters	
β_1		β_{avg}	b	d
0.5	(Benchmark)	0.8186	28.2935	0.103
0.35		0.8186	16.0373	0.0301
0.65		0.8186	-224.071	-1.75
0.5		0.8983	10.2472	0.00383
0.35		0.8983	7.5292	0.00085
0.65		0.8983	17.1918	0.0202
0.5		0.703	-43.2913	-0.73
0.35		0.703	94.3404	0.93
0.65		0.703	-7.3796	-0.3504

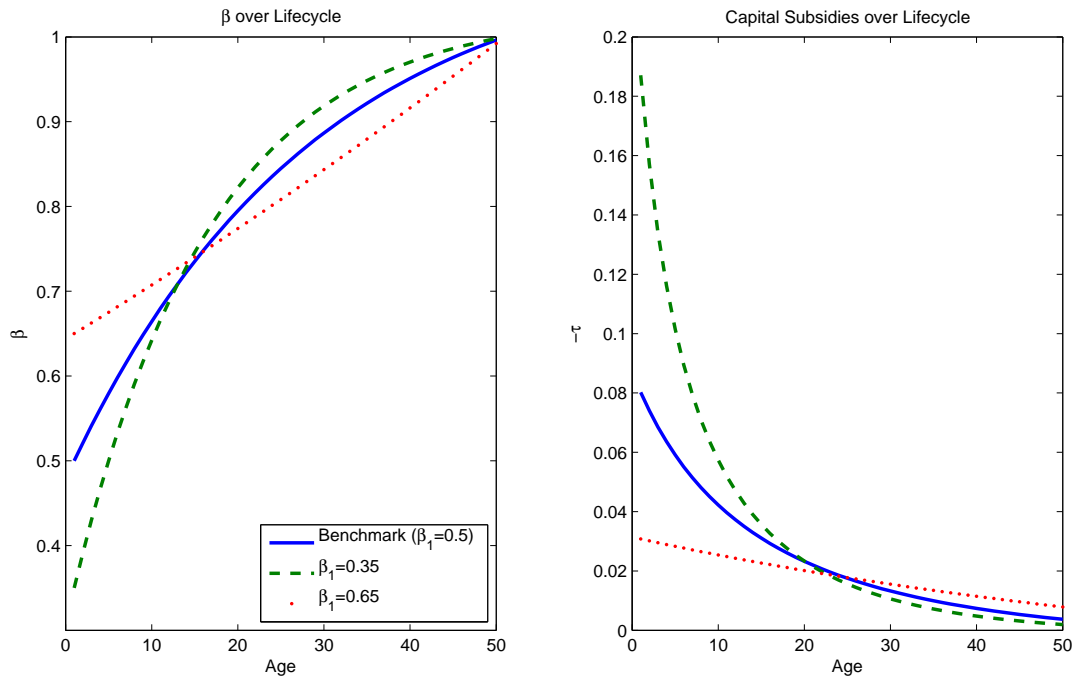


Figure 3: Optimal subsidies for benchmark calibration and sensitivity with respect to β_1 .

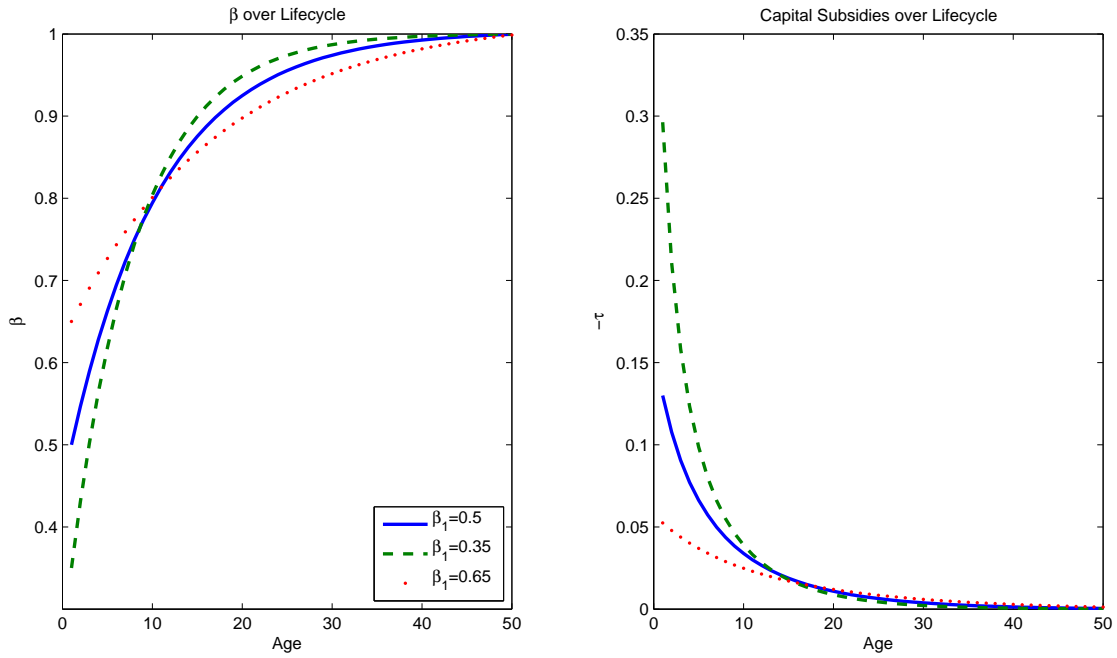


Figure 4: Optimal subsidies for calibration with higher β_{avg} and sensitivity with respect to β_1 .

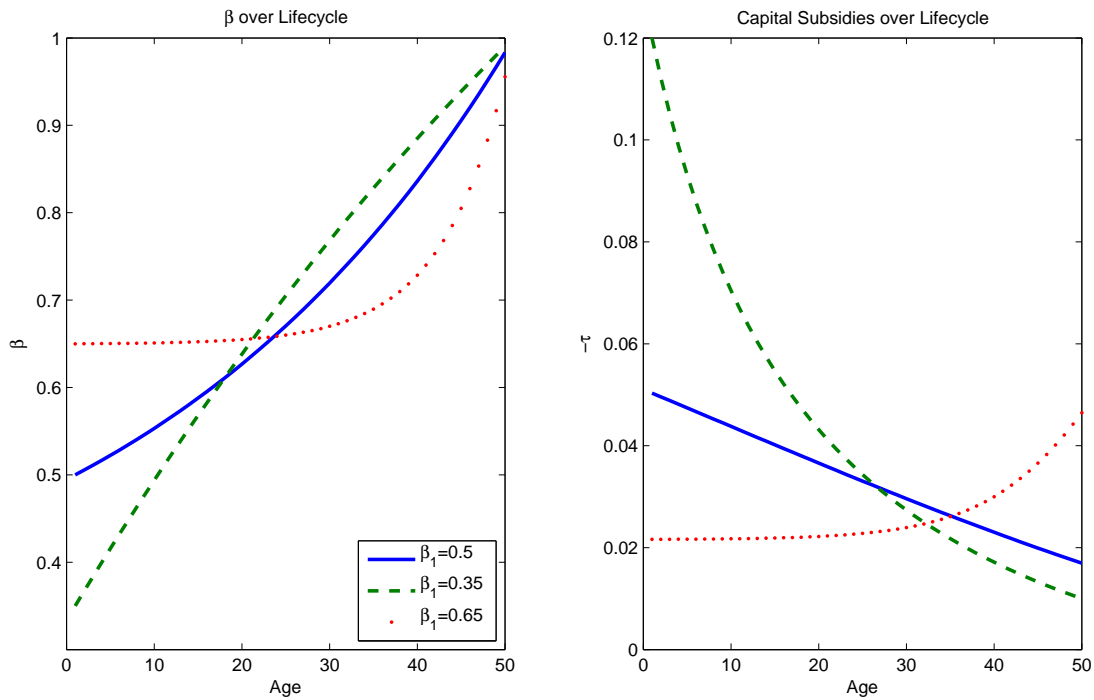


Figure 5: Optimal subsidies for calibration with lower β_{avg} and sensitivity with respect to β_1 .

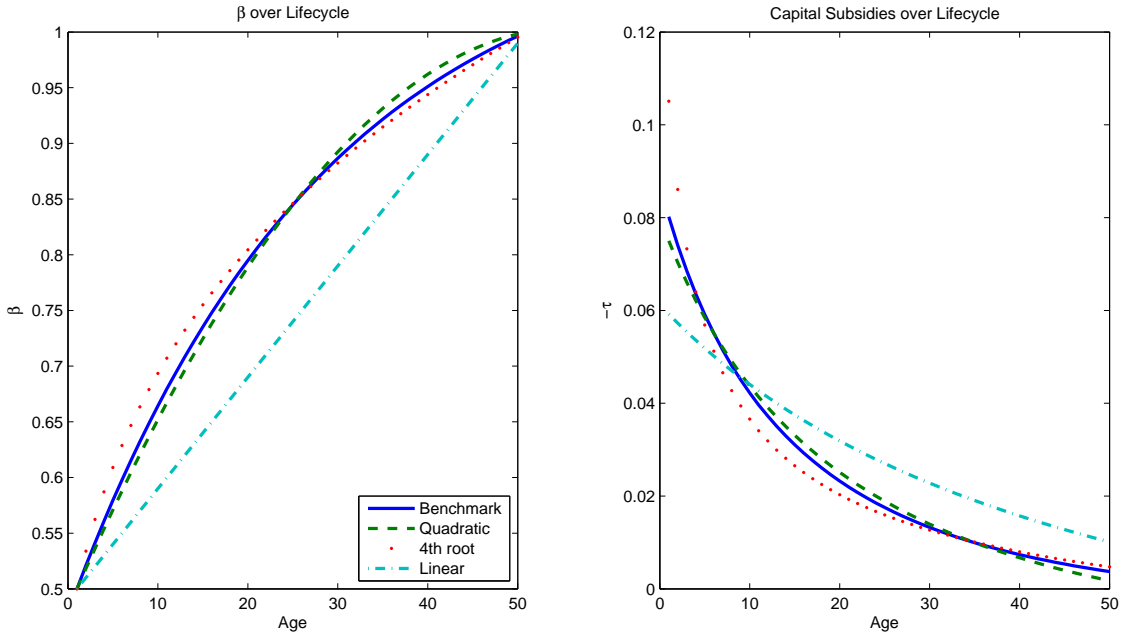


Figure 6: Benchmark case and sensitivity with respect to functional forms and curvature of β_i .

of self-control problems to a higher level of β_1 , namely 0.65, this affects optimal subsidies in two ways. First, the level of subsidies in the early periods of the life cycle are much lower: instead of starting at 8% as in the benchmark case, they start at below 4%, and benchmark subsidies remain significantly larger than $\beta_1 = 0.65$ case in the early years of the life cycle. This is mainly due to the fact that, in the $\beta_1 = 0.65$ case, younger people begin life with less self-control problems, and thus, the current component of the optimal tax formula calls for lower levels of subsidies compared to the benchmark case. Second, when $\beta_1 = 0.65$, the decline of subsidies with age is milder relative to the benchmark case. This is due to the fact that, for a given level of β_{avg} target, a higher β_1 target calibrates a β_i function that is less concave relative to the benchmark case, which as Corollary 5 suggests, implies a milder decline in subsidies with age. In fact, when $\beta_1 = 0.65$, the self-control problem as a function of age is slightly convex, but not convex enough to make optimal subsidies increase with age. The discussion of the optimal subsidies with a calibration target of $\beta_1 = 0.35$ is analogous and hence is omitted.

In Figure 4 and Figure 5, we analyze sensitivity with respect to calibration target β_{avg} . The calibrated parameters of the $i \rightarrow \beta_i$ function are reported in rows four to nine of Table 1. Comparison of the blue lines in Figure 3 and Figure 4 shows that, for a given level of β_1 target, a higher β_{avg} target, namely 0.898 relative to the 0.818 in the benchmark case, gives a calibrated self-control function that is more concave in age. This results in higher optimal subsidies early in the life cycle and sharper decline in subsidies with age. A comparison of blue lines in Figure 3 and Figure 5, where in this case we consider a β_{avg} target that is lower than the benchmark target, namely 0.703, confirms our findings regarding how β_{avg} affects optimal subsidies over the life cycle. Figure 5 is also interesting because we see that when calibration targets for β_1 and β_{avg} are close to each other, the calibration procedure gives a $i \rightarrow \beta_i$ function that is convex enough to make optimal subsidies increase with age.

In Figure 6, we provide sensitivity analysis with respect to our choice of functional form for the $i \rightarrow \beta_i$ function. The figure shows that, for benchmark calibration targets, if we instead choose a quadratic or 4th order polynomial functional form, the resulting optimal subsidies over the life cycle follows a pattern that is very similar to our benchmark case.

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

How Large Are the Optimal Subsidies Relative to Existing Ones? Observe that in our computations the tax base is the gross return on asset holdings. Most actual tax systems, however, tax *asset income*. If we translate our numbers taking that into account, we find that optimal subsidies on capital income at the earliest age takes a value of approximately 200%, decreasing to 9%.²⁵ These are obviously large numbers. In this section, we compare these numbers with existing saving subsidies in the United States.

Tax deferred retirement accounts. As the name suggests, a tax deferred retirement account

²⁵Denoting 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 $-\tau_1 = .08$ and $R = 1.04$ we have $-\tau_1^k = 2$.

allows to defer tax payment till retirement. It works through *deductible contributions*. Probably, the most common defined contribution plan linked to a tax deferred account in the United States is the classical 401(k). Each dollar invested into a classical 401(k) can be deducted from taxable income. The 401(k) restricts withdrawals before retirement and pays taxes both on the principal and the interest generated only at the date of withdrawal, according to the tax rate faced by the agent at that date. Moreover, it is quite likely that, at retirement age, contributors face lower marginal taxes than when they invested into the plan. As we see below, this feature may generate considerable saving subsidies, and, importantly, these saving subsidies depend on where the agent is over his life cycle. Consider an agent who is at age i and who is facing a marginal income tax rate τ_C^y based on the income tax bracket she falls in. Suppose there are N periods before she retires. If she invests \$1 today in 401(k), with the current tax deduction, this is as if she invests $\frac{1}{1-\tau_C^y}$. If τ_R^y is the income marginal tax rate at retirement age, the agent will receive $\$R^N \frac{1-\tau_R^y}{1-\tau_C^y}$ at retirement. Therefore, the yearly gross return for each dollar invested in 401(k) at age i - and N periods before retirement - is given by $\left(R^N \frac{1-\tau_R^y}{1-\tau_C^y}\right)^{\frac{1}{N}}$, which implies a yearly saving subsidy of

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

This saving subsidy depends on a person's age in two ways.²⁶ First, a person's age determines how far away he is from retirement, N , which clearly affects the subsidy rate in (6). Second, people's income tend to depend where they are on their life cycle, which implies the tax deduction they receive, τ_C^y , effectively depends on their age. Observe that in computing the implied 401(k) subsidy rate in (6) we take the tax base as the gross return on asset holdings to make it comparable to the optimal subsidies we have computed earlier in this section.

401(k) plan with employer matching. The 401(k) also allows the employer to contribute to the

²⁶The formula in (6) indicates that the implied subsidy rate of the 401(k) scheme is independent of the agent's saving rate. The 401(k) scheme, however, puts a cap on the amount agents can invest. In 2012, the maximal amount for agents aged 50 or below was \$17,000; older contributors faced an higher cap. As we will explain below however, the cap is unlikely to be binding for the median household with an average saving rate.

worker investment plan. The most common methods of employer matching are the \$1 per \$1 up to 6% of pay and the \$0.5 per \$1 up to 6% of pay.²⁷ According to these options, if the saving rate is below 6%, then for each dollar that a worker contributes to the 401(k) account, the employer contributes one (or, respectively, one half) dollar. This means a worker investing one dollar is effectively investing $\frac{\$2}{(1-\tau_C^y)}$ into the plan when the employer matching is 1-to-1 and $\frac{\$1.5}{(1-\tau_C^y)}$ when the employer matching is 0.5-to-1. The formula in (6) can then straightforwardly be adapted to compute implied 401(k) subsidies in the presence of employer matching.²⁸

Now, we compare the saving subsidies implied by a typical 401(k) plan to the optimal subsidies implied by our model. In Figure 7, right panel, we report the life-cycle profile of the median income per household head - between 17 and 67 years of age - in the period 2000-2006 and the corresponding marginal tax rates implied by the 2006 income tax code.²⁹ In Figure 7, left panel, we report the implied saving subsidies for several 401(k) plans together with the optimal subsidies given by our model under the benchmark parameterization. Three observations are immediate. First, interestingly, the range of values for the subsidies implied by the 401(k) plan are not very different from the optimal ones. Second, the subsidies implied by the 401(k)

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

²⁸Let a person's annual income be w and his amount contributed to 401(k) be x . If $x > 0.06w$, meaning the person is contributing more than 6% of his income, then in a one-to-one matching plan, the employer contributes $0.06w$ dollars, which implies that for each dollar he invests he is effectively investing

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

Therefore, in this case, the implied subsidy depends on the amount contributed. However, Thaler and Benartzi (2004) report that the average saving rates into the SMarT plan (for the 'control group') is between 4.4% and 6.6% (see page S174). Thus, in our computations of the implied 401(k) subsidies, we assume that contribution rate is less than 6% and use the formula explained in the main text.

²⁹The data for the life-cycle profile of the median income per household head in the period 2000-2006 is taken from Heathcote, Perri, and Violante (2010).

plan are very much age-dependent. Third, the life-cycle pattern of the 401(k) subsidies is qualitatively very different from the optimal ones as they are *increasing over the life-cycle*. Existing subsidies appear too low for young individuals and too high for individuals close to retirement.

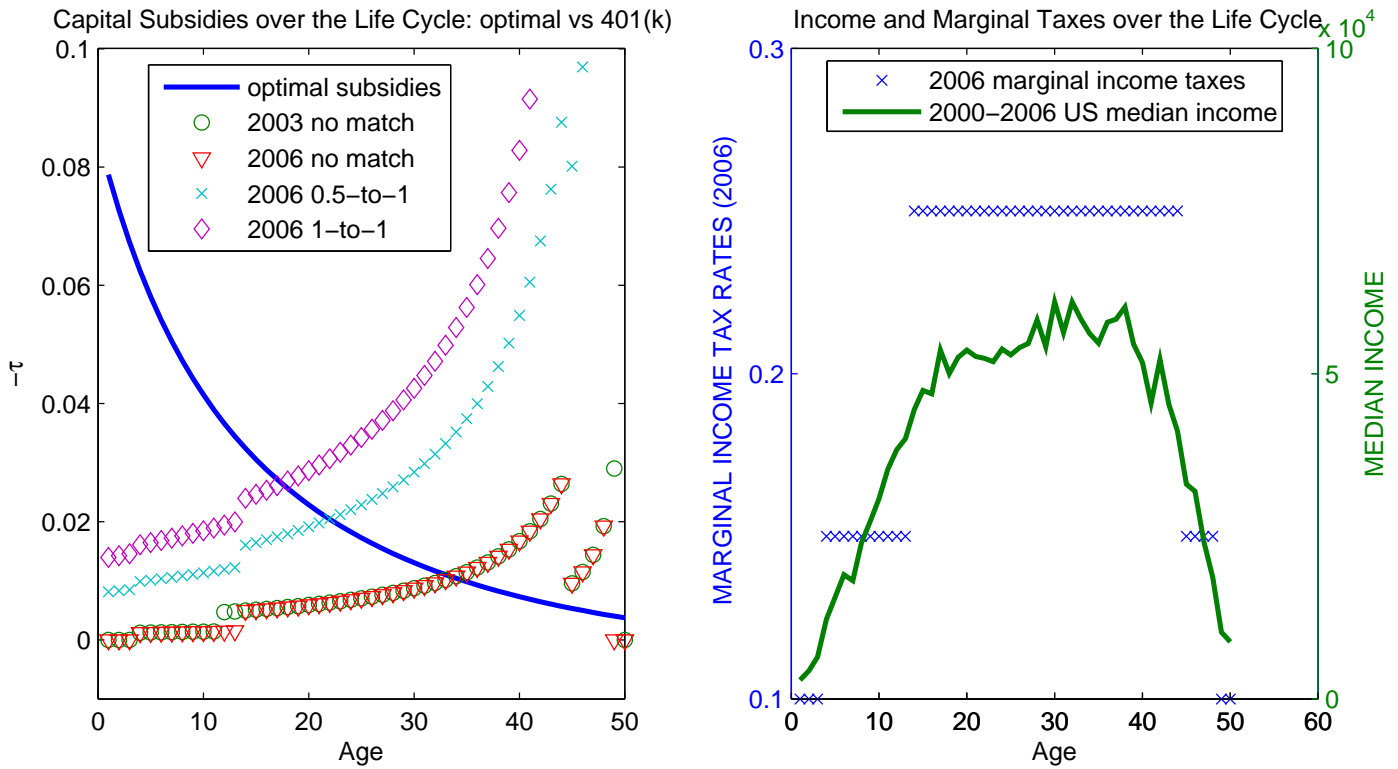


Figure 7: Left panel: Saving subsidies implied by the 401(k) plan according to the marginal income tax rates in 2003 and 2006, at different levels of employer matching. Right panel: Median income of the U.S. household head over the life cycle in 2000-2006 and implied marginal income taxes in the year 2006.

Comparison to the Constant Self-Control Model. How the level of optimal capital subsidies we obtain in our experiments compare to those one would obtain in a model where self-control problems are constant over the life cycle? Table 2 reports optimal tax levels implied by models with constant self-control problems for four representative levels of β . The first column of the

Table indicates the level of self-control used in the computations. For example, if β is constant at the value estimated by Laibson, Repetto, and Tobacman (2007), namely 0.818, optimal subsidies on capital are 0.93% independent of age. On the other hand, in the benchmark simulation depicted in Figure 3 by the solid blue line, optimal capital subsidies are much higher than 0.93% for a large fraction of the life cycle. The optimal subsidy is as high as 8% at the beginning of the life cycle. This might not be very surprising since people have a lower level of self-control at earlier ages in our model, $\beta_1 = 0.5$, relative to the constant self-control model where $\beta = 0.818$ at every age. What is perhaps more surprising is that the level of optimal subsidies remain higher than the one implied by the constant self-control model until the very end of the life cycle. For instance, in period 22, the agent in our model has the same degree of self-control as in the constant self-control model, i.e., $\beta_{22} \approx 0.818$. However, the optimal subsidy he receives is 2.08%, which is much larger than 0.93%.

To see how the level of optimal subsidies in our model compares to the constant self-control model, one can rewrite the optimal tax formula given by Proposition 1 as

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

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

To grasp the intuition why our model implies higher subsidies, remember the decomposition of optimal taxes into the *current* and the *future* components from the discussion following Proposition 2. The current component is related to an agent's current degree of self-control problem and calls for undersaving. For agent in period 22, this component is the same between our model and the constant self-control model. The future component summarizes how

Table 2: Optimal Subsidies under Constant Self-Control Problems

β	0.818	0.65	0.5	0.35
$-\tau$	0.93%	2.15%	4%	7.43%

much a person oversaves to compensate for future self's undersaving. Since people's degree of self-control improves with age in our model, the future component makes the agent save more today in the constant self-control model relative to ours. As a result, an agent with the same level of current self-control problem saves more in the constant self-control model, which implies the required subsidy to make him save the right amount is going to be lower. In periods after 22, people in our model have lower degrees of self-control problems relative to the constant self-control model with an average $\beta = 0.818$, but their future selves have even lower degrees of self-control problems. Therefore, the current component of the optimal tax formula calls for subsidizing them less relative to the agents in the constant self-control model, but the future component calls for the opposite. Figure 3 suggests that optimal subsidies are higher in our model compared to constant self-control until well beyond period 22. More precisely, in the benchmark calibration, the solid blue line stays above 0.93 till period 37 (i.e., until age 57).

4 Effect of Imperfect Altruism and Partial Sophistication

We have seen above that taking into account the life-cycle patterns of self-control problems has a significant quantitative effect on the level of optimal capital subsidies. Due to the assumption of logarithmic utility or steady-state, agents' level of sophistication has not mattered for our results. Furthermore, we have assumed perfect altruism across generations. In this section, we remove each of these assumptions one by one and investigate the quantitative importance of imperfect altruism and partial sophistication on optimal subsidies. The main conclusion is that they both have little effect on both the level and pattern of optimal saving subsidies.

4.1 Imperfect Altruism

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

A parent has the following preferences over dynastic consumption streams

$$u(c_0) + \gamma \left[\delta u(c_1) + \delta^2 u(c_2) + \dots + \delta^I u(c_I) + \delta^{I+1} u(c'_0) + \delta^{I+1} \gamma \left[\delta u(c'_1) + \delta^2 u(c'_2) + \dots \right] \dots \right],$$

where this preference specification is equivalent to the one in the baseline model whenever the altruism factor, γ , is equal to 1. When $\gamma \in [0, 1)$, there is imperfect altruism. The finite life time case of Krusell, Kuruscu, and Smith (2010) corresponds to the case of $\gamma = 0$. The rest of the parent's problem is identical to the one in Section 2.3.

Proposition 6 generalizes optimal tax formulas of Proposition 1 to the case with a general altruism factor, γ , under the assumption of full sophistication at all ages.

Proposition 6 *Suppose $u(c) = \log(c)$ and people are fully sophisticated. Then, for $\gamma \in (0, 1]$ we have:*

$$1 - \tau_0^* = \frac{1 + \beta_1 \delta [1 + \delta + \dots + \delta^{I-2} + \delta^{I-1} D]}{1 + \delta + \dots + \delta^{I-1} + \delta^I D},$$

$$1 - \tau_i^* = \frac{1}{\beta_i} \frac{1 + \beta_{i+1} \delta [1 + \delta + \dots + \delta^{I-i-2} + \delta^{I-i-1} D]}{1 + \delta + \dots + \delta^{I-i-1} + \delta^{I-i} D}, \text{ for } i \in \{1, \dots, I-1\},$$

$$1 - \tau_I^* = \frac{1}{\beta_I},$$

where

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

Proof. Relegated to Appendix A. ■

It is straightforward to see that in the case of perfect altruism, $\gamma = 1$, these formulas reduce back to the ones in Proposition 1.

In Figure 8, we analyze the quantitative effect of the finite life time channel on the monotonicity properties of optimal subsidies. The solid blue line represents optimal subsidies under the benchmark calibration with $\gamma = 1$, whereas the dashed blue line represents those under benchmark calibration targets but with $\gamma = 0$. The comparison of the solid and dashed lines show that the finite life time effect is not strong enough to overturn the optimality of subsidies declining with age. To study the robustness of the decreasing subsidy result, in the figure, we also report the results for a β_i function that is increasing linearly with age starting from $\beta_1 = 0.5$. This shape is even less concave than our benchmark calibration, and it corresponds to the linear dashed line of Figure 6. In this case too, a comparison of the solid red line in Figure 8, which represents the perfect altruism case, and the dashed red line, which represents the no altruism case, show that the finite life time effect is not strong enough to invalidate the optimality of declining subsidies. Observe that in both robustness exercises we set $\gamma = 0$. For $\gamma \in (0, 1)$, the effect of finite life time on the monotonicity properties of optimal subsidies would be even smaller. We simulated our model adopting several different parameterizations of γ, δ and $i \rightarrow \beta_i$. The optimality of decreasing saving subsidies is a quite robust finding (details are available upon request). What seems to play an important role to maintain robustness of the decreasing subsidy result under various values of γ is that, in our model, self-control problems vanish towards the end of the life cycle.

We conclude this section by providing intuition for why in Figure 8, the subsidy rates at all ages are uniformly higher in the model with imperfect altruism (dotted lines) compared to our benchmark model with perfect altruism (solid lines). Mathematically, observe that the parameter γ enters into the formula for $1 - \tau_i^*$ in Proposition 6 only through the constant D . It is easy to show that D is increasing in γ and $1 - \tau_i^*$ is decreasing in D . To grasp the intuition, recall our decomposition of the optimal tax formula into current and future components, discussed right after Proposition 2. In light of that decomposition, the first term on the right-hand side of the formula for $1 - \tau_i^*$ in Proposition 6, $\frac{1}{\beta_i}$, is the current component and always calls for a subsidy. The second term, which is where the altruism factor enters the tax formula, is the future component and always calls for a tax. The future component is there because agent i

at date t disagrees with agent $i + 1$ regarding how agent $i + 1$ should discount (utility from) consumption in period $t + 2$ and onwards relative to consumption in $t + 1$: from agent i 's perspective, the correct discount factor between consumption in date $t + 1$ and that in $t + s$ is δ^{s-1} , whereas agent $i + 1$ discounts the same utility from consumption by $\beta_{i+1}\delta^{s-1}$. To correct for the eventual undeserving of self $i + 1$, self i oversaves relative to the efficient allocation, and in order to prevent this, the government has to tax self i . If $\gamma = 0$, the disagreement between self i and self $i + 1$ regarding the discounting between $t + 1$ consumption and future consumption levels stop at the end of the current life cycle ($D = 1$), while for $\gamma = 1$, the disagreement piles up for infinitely many generations ($D = \frac{1}{1-\delta}$). As a result, when $\gamma = 0$, there is less cumulative disagreement between selves i and $i + 1$, which means self i 's oversaving relative to the efficient allocation is lower, which implies the tax implied by future component is lower. Therefore, the future component mitigates the subsidy implied by current component to a lesser degree, and hence, the subsidy is larger.

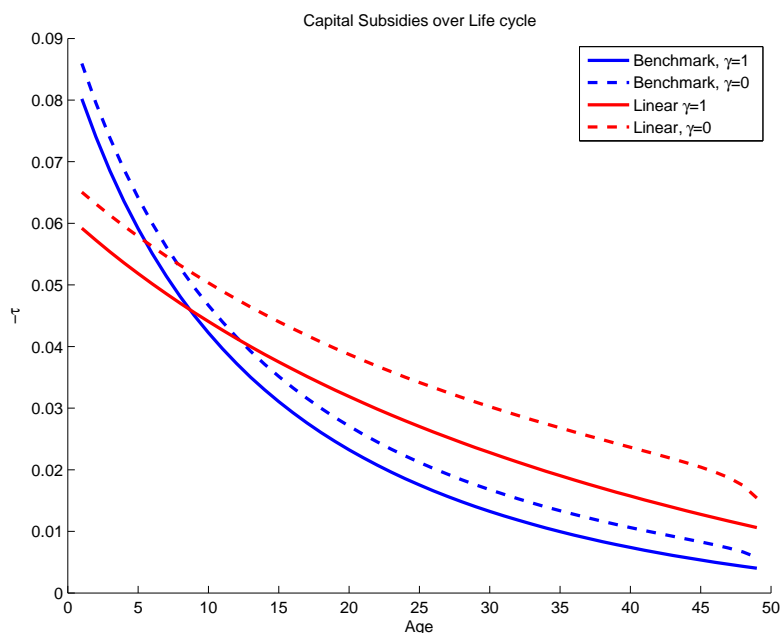


Figure 8: Sensitivity of life-cycle subsidies to the degree of parent's altruism γ .

4.2 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. The details of our computational procedure are explained in Appendix B.

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. Throughout this section, we set the pattern $i \rightarrow \beta_i$ according to our benchmark calibration, i.e. the first line of Table 1. In Figure 9, 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 9, 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 subsidies when the level of sophistication changes smoothly over the life cycle. We assume π increases concavely. This experiment is summarized in Figure 10, 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 10 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 10 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 (decrease) with the level of sophistication. It can be formally shown that this pattern is quite general.³⁰

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 life cycle 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 life cycle is not significantly affected by our choice of the coefficient of constant relative risk

³⁰An earlier related result is given in O'Donoghue and Rabin (2003a) 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 (2003a) 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. It can be shown that this result is valid under our way of modeling partial sophistication as well (the derivations are available upon request).

aversion.

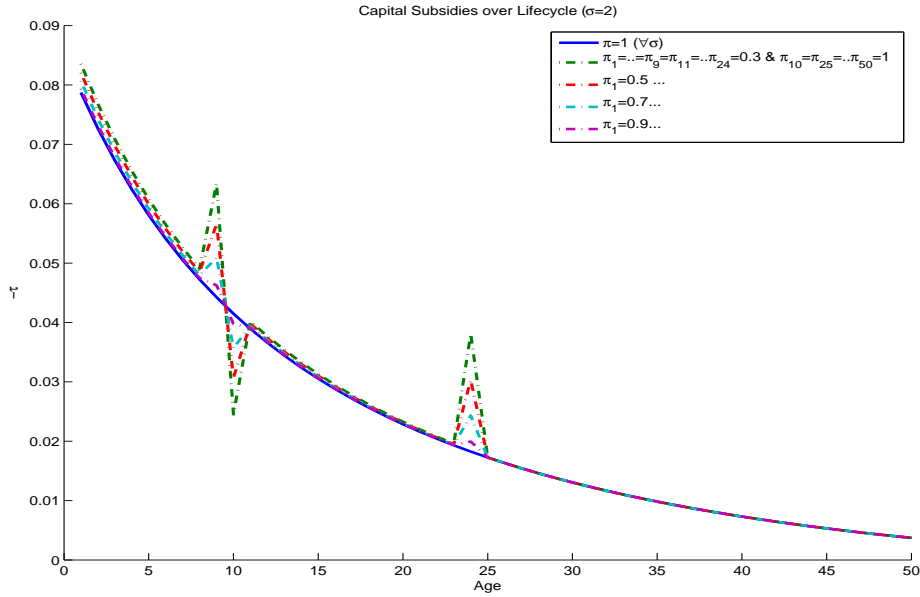


Figure 9: Sensitivity with respect to partial sophistication (jumps in π)

5 Conclusion

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

Capital subsidies should start somewhere between 3% and 18% at the beginning of the life cycle and decline monotonically with age to somewhere between 0% and 1%, depending on

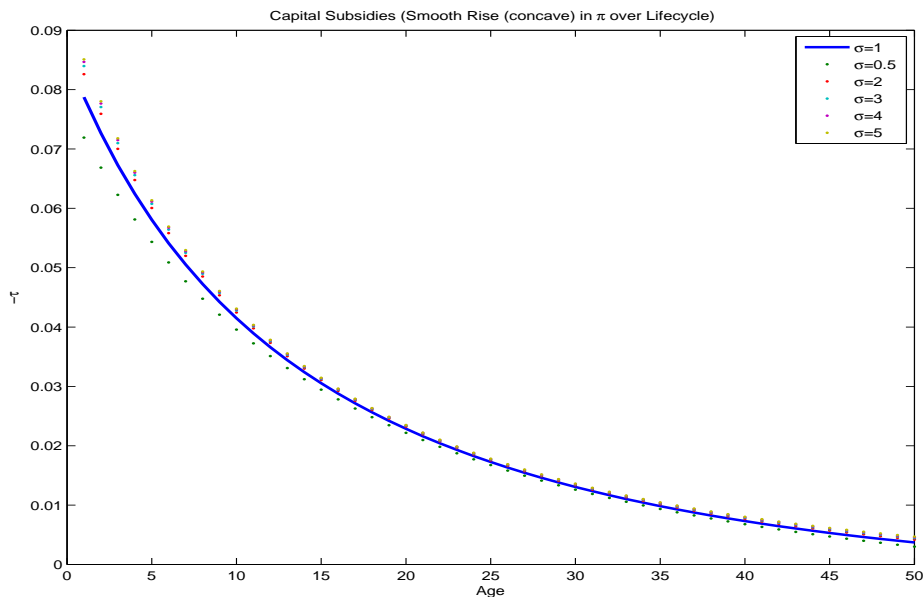


Figure 10: *Sensitivity with respect to partial sophistication (smoothly rising π)*

the particular parameterization of the model. If we translate them into subsidies to capital income, these are very large numbers. Perhaps more importantly, we show they are much larger than the savings subsidy we would obtain in models with constant self-control, at almost all ages. Our model is probably too simple for delivering precise policy predictions. Nevertheless, our analysis suggests that researchers who take self-control problems seriously should also carefully measure the evolution of self-control problems over the life cycle before making policy suggestions.

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

The existence of illiquid assets does not change our optimal tax results as long as there are

no borrowing constraints. More precisely, in Appendix D we use a three periods example to show that a tax system that is optimal in an environment without illiquid assets is still optimal in the same environment with an illiquid asset as long as we complement the tax system with an appropriate tax on the illiquid asset.

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.
- ASHEIM, G. B. (2007): "Procrastination, Partial Naivete, and Behavioral Welfare Analysis," Memorandum 02/2007, Oslo University, Department of Economics.
- 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), 315–372.
- DELLAVIGNA, S., AND U. MALMENDIER (2004): "Contract Design and Self-control: Theory and Evidence," *The Quarterly Journal of Economics*, 119(2), 353–402.
- ELIAZ, K., AND R. SPIEGLER (2006): "Contracting with Diversely Naive Agents," *Review of Economic Studies*, 73(3), 689–714.

- EROSA, A., AND M. GERVAIS (2002): "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 105(2), 338–369.
- FARHI, E., AND I. WERNING (2013): "Insurance and Taxation over the Life Cycle," *Review of Economic Studies*, 810(2), 596–635.
- 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.
- GOLOSOV, M., M. TROSHKIN, AND A. TSYVINSKI (2011): "Optimal Dynamic Taxes," NBER Working Papers 17642, National Bureau of Economic Research, Inc.
- 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(1), 89–96.
- GRUBER, J., AND B. KOSZEGI (2004): "Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes," *Journal of Public Economics*, 88(9-10), 1959–1987.
- GUL, F., AND W. PESENDORFER (2001): "Temptation and Self-Control," *Econometrica*, 69(6), 1403–1435.
- HARRIS, C. J., AND D. LAIBSON (2001): "Dynamic Choices of Hyperbolic Consumers," *Econometrica*, 69(4), 935–957.
- HEATHCOTE, J., F. PERRI, AND G. L. VIOLANTE (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, 13(1), 15–51.

- 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.
- IMROHOROGLU, A., S. IMROHOROGLU, AND D. H. JOINES (2003): "Time-Inconsistent Preferences And Social Security," *The Quarterly Journal of Economics*, 118(2), 745–784.
- 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.
- 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," *Ph.D. Dissertation*.
- (1997): "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112(2), 443–477.
- LAIBSON, D., A. REPETTO, AND J. TOBACMAN (2007): "Estimating Discount Functions with Consumption Choices over the Lifecycle," NBER Working Paper 13314, National Bureau of Economic Research.
- 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.

- (2003a): “Self Awareness and Self Control,” in *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, pp. 217–243. Russell Sage Foundation Press.
- (2003b): “Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes,” *American Economic Review*, 93(2), 186–191.
- (2006): “Optimal Sin Taxes,” *Journal of Public Economics*, 90(10-11), 1825–1849.
- PAVONI, N., AND H. YAZICI (2012): “Optimal Life-cycle Capital Taxation under Self-Control Problems,” Icier working paper 467.
- (2013): “Present-Bias Preferences and Optimal Taxation of Parental Transfers,” mimeo.
- PHELPS, E. S., AND R. A. POLLAK (1968): “On Second-Best National Saving and Game-Equilibrium Growth,” *The Review of Economic Studies*, 35(2), 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 S. BENARTZI (2004): “Save More Tomorrow (TM): Using Behavioral Economics to Increase Employee Saving,” *Journal of Political Economy*, 112(S1), S164–S187.
- THALER, R. H., AND H. M. SHEFRIN (1981): “An Economic Theory of Self-Control,” *Journal of Political Economy*, 89(2), 392–406.

A Proofs

A.1 Proof of Proposition 1.

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

Preparation to the proof.

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

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

Agents face a price sequence satisfying:

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

that is, it is generated by a capital stock sequence $\{k_t^*\}_t$ where the capital stock is generated by K . Since the problem is recursive, a government which aims to implement the efficient allocation will use the same taxes in any two periods if the age of the agent and the capital stock in those periods are the same. Therefore, without loss of generality, we define taxes as functions of age and capital stock as follows: $\tau_i(k_t)$ is the savings (capital) tax agent at age $i = 0, 1, \dots, I$ 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 $b_i(b_{i-1}(\dots b_{n+1}(b_n, k_t^*; \tau) \dots), k_{t+i-n-1}^*; \tau), k_{t+i-n}^*; \tau)$, where for all t, s , we define $k_{t+s}^* = K(K(\dots(k_t^*) \dots))$, where the K function has been applied s times. To simplify notation, we will denote this mapping simply as $b_i(\dots(b_n) \dots)$.

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: $k^i = K(K(\dots(k) \dots))$, where the function K has been applied i times. In the problem below, the function K is fixed to that

of the commitment allocation. Of course, the function describing the evolution of aggregate capital in equilibrium is part of the fixed point argument as it must satisfy market clearing.

Parent's Problem along the Transition

$$V(b, k; \tau) = \max_{b_0} u(R(k)(1 - \tau_1)b + w(k) + T_1 - b_0) + \\ + \delta \left[\sum_{i=0}^{I-1} \delta^i u \left(R(k^{i+1})(1 - \tau_i) b_i(\dots(b_0)\dots) + w(k^{i+1}) + T_i - b_{i+1} \right) + \delta^I V \left(b_I(\dots(b_0)\dots), k^{I+1}; \tau \right) \right] \\ \text{s.t. for all } b_0$$

$$b_1(b_0, k^1; \tau) = \arg \max_{\hat{b}_1} u \left(R(k^1)(1 - \tau_0)b_0 + w(k^1) + T_0 - \hat{b}_1 \right) + \\ + \delta \beta_1 \left[\pi_1 \left\{ \sum_{i=1}^{I-1} \delta^{i-1} u \left(R(k^{i+1})(1 - \tau_i) b_i(\dots(\hat{b}_1)\dots) + w(k^{i+1}) + T_i - b_{i+1}(\dots(\hat{b}_1)\dots) \right) + \delta^{I-1} V \left(b_I(\dots(\hat{b}_1)\dots), k^{I+1}; \tau \right) \right\} \right. \\ \left. + (1 - \pi_1) W_1 \left(\hat{b}_1, k^2; \tau \right) \right] \\ \text{s.t. for all } b_1$$

$$b_2(b_1, k^2; \tau) = \arg \max_{\hat{b}_2} u \left(R(k^2)(1 - \tau_1)b_1 + w(k^2) + T_1 - \hat{b}_2 \right) \\ + \delta \beta_2 \left[\pi_2 \left\{ \sum_{i=2}^{I-1} \delta^{i-2} u \left(R(k^{i+1})(1 - \tau_i) b_i(\dots(\hat{b}_2)\dots) + w(k^{i+1}) + T_i - b_{i+1}(\dots(\hat{b}_2)\dots) \right) + \delta^{I-2} V \left(b_I(\dots(\hat{b}_2)\dots), k^{I+1}; \tau \right) \right\} \right. \\ \left. + (1 - \pi_2) W_2 \left(\hat{b}_2, k^3; \tau \right) \right] \\ \text{s.t. for all } b_2$$

...

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

$$b_{I-1}(b_{I-2}, k^{I-1}; \tau) \in \arg \max_{\hat{b}_{I-1}} u \left(R(k^{I-1})(1 - \tau_{I-2})b_{I-2} + w(k^{I-1}) + T_{I-2} - \hat{b}_{I-1} \right) + \delta \beta_{I-1} (1 - \pi_{I-1}) W_{I-1} \left(\hat{b}_{I-1}, k^I; \tau \right) \\ + \delta \beta_{I-1} \left[\pi_{I-1} \left\{ u \left(R(k^I)(1 - \tau_{I-1})\hat{b}_{I-1} + w(k^I) + T_{I-1} - b_I(\dots(\hat{b}_{I-1})\dots) \right) + \delta V \left(b_I(\dots(\hat{b}_{I-1})\dots), k^{I+1}; \tau \right) \right\} \right] \\ \text{s.t. for all } b_{I-1}$$

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

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

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

with

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

Letting b_i and k^{i+1} be the saving level in period i and aggregate capital stock in period $i + 1$, define (we

disregard the tax dependence for notational simplicity):

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

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

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

Note that, using

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

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

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

and

$$\Gamma_0(b_0(b, k; \tau), k^1) = R(k^1)(1 - \tau_0(k^1))(1 - M_0)\Gamma_I(b, k),$$

where

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

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

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

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

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

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

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:

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

Core proof of Proposition 1.

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

Guess

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

where D and $D_0, D_1, \dots, D_I, B_0, \dots, B_I$ 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, \dots, D_I .

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

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

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

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

and

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

Thus,

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

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

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

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

where we used the recursion (8).

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

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

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 $i = 1, \dots, I - 1$:

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

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 :

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

which implies

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

and hence

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

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

$$M_i = \frac{1 - \delta}{1 - \delta + \beta_i \delta}, \text{ for all } i \in \{1, \dots, I\}, \quad (11)$$

$$M_0 = 1 - \delta.$$

Now we turn to taxes that implement the efficient allocation. The constraint that describes self- i 's behavior for $i \in \{1, \dots, I - 1\}$ becomes the following once we plug in the derivatives of the value functions from (16) :

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

The comparison of (12) with the efficiency condition (1) gives the optimal tax as:

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

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

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

and the comparison of this with the efficiency condition gives

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

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

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

with the corresponding optimality condition gives

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

A.2 Proof of Proposition 2.

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

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

which renders optimal taxes as:

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

Under CEIS utility and linear policies, we have:

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

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

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

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

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

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

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

Plugging (13) in (14), we get a system of $(I + 1)$ equations in $(I + 1)$ unknowns (M_0^*, \dots, M_I^*) that fully pin down agents policies when they face optimal taxes, for the CEIS case:

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

Clearly, the solution to this system does not depend on σ . In fact, it is easy to show that the logarithmic utility solution given by equation (11) satisfies the above system of equations, meaning it is an equilibrium. Plugging (11) in the formula for taxes, (13), we get that optimal taxes are the same as the logarithmic utility case.

A.3 Proof of Proposition 6.

The proof of Proposition 6 follows the proof of Proposition 1 very closely. The important difference is that the altruism factor, γ , can be any number in $[0, 1]$. In this case, the maximization problem of the parent is identical to (9), except that the objective function has the general altruism factor:

$$V(b, k; \tau) = \max_{M_0} u(M_0 \Gamma_I(b)) + \gamma \delta \left[\sum_{i=1}^{I-1} \delta^i u(Q_{i-1}(k) M_i \Gamma_I(b)) + \delta^I V\left((1 - M_I) Q_{I-1}(k) \Gamma_I(b), k^{I+1}; \tau\right) \right]. \quad (15)$$

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

Note that since we assume full sophistication, meaning $\pi_i = 1$ for all i , the naive value function do not appear in the planner's problem. Therefore, we only guess and verify parental value function.

Guess

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

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

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

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

where we used the recursion (8).

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

$$(M_i Q_{i-1}(k))^{-1} = \delta \beta_i R(k^{i+1})(1 - \tau_i(k^{i+1})) (Q_i(k))^{-1} \left[\sum_{j=i+1}^I \delta^{j-(i+1)} + \delta^{I-i} D \right]$$

and

$$(M_I Q_{I-1}(k))^{-1} = \delta \beta_I R(1 - \tau_I(k^{I+1})) (Q_I(k))^{-1} D.$$

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

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

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

$$M_i(D) = \frac{1}{1 + \beta_i \delta \left\{ \sum_{j=i+1}^I \delta^{j-(i+1)} + \delta^{I-i} D \right\}}. \quad (16)$$

Taking first-order condition with respect to bequests in the parent's problem (9) and plugging in the $M_i(D)$ from above for all i , we get:

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

Now verify the value function to compute D :

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

which implies

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

and hence

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

Now we turn to taxes that implement the efficient allocation. The constraint that describes self- i 's behavior for $i \in \{1, \dots, I-1\}$ becomes the following once we plug in the derivatives of the value functions from (16) :

$$(M_i Q_{i-1}(k) \Gamma_I(b, k))^{-1} = \delta \beta_i R(k^{i+1}) (1 - \tau_i(k^{i+1})) (M_{i+1} Q_i(k) \Gamma_I(b, k))^{-1} \left[\sum_{j=i+1}^I \delta^{j-(i+1)} + \delta^{I-i} D \right] M_{i+1}. \quad (17)$$

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

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

which, using (16), implies

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

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

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

and the comparison of this with the efficiency condition gives

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

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

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

with the corresponding optimality condition gives

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

B Computational Procedure

B.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 $i = 0, 1, \dots, I$ 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.

B.2 Characterizing equilibrium value function constants for a given tax system τ :

STEP 1: Computing equilibrium D_0, \dots, D_I .

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

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

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

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

which together imply

$$D = M_0^{-\sigma}. \quad (19)$$

B.3 Characterizing optimal tax system, τ^* :

The incentive constraints for agents $i = 1, \dots, I$ 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 commitment Euler equations, we immediately see that optimal taxes must satisfy:

$$\begin{aligned} \text{For all } i &\in \{0, \dots, I - 2\}, & (20) \\ (1 - \tau_{i+1}^*) &= \frac{1}{\beta_{i+1}} \left(\frac{\left[\pi_{i+1} \left\{ \sum_{j=i+2}^I \delta^{i-(j+2)} \left(M_j^* \frac{Q_{j-1}^*}{Q_{i+1}^*} \right)^{1-\sigma} + \delta^{I-(i+1)} D^* \left(\frac{Q_i^*}{Q_{i+1}^*} \right)^{1-\sigma} + (1 - \pi_{i+1}) D_{i+1}^* \right\} \right]}{M_{i+2}^{*-\sigma}} \right)^{-1} \\ (1 - \tau_I^*) &= \frac{1}{\beta_I} \left(\frac{[\pi_I D^* + (1 - \pi_I) D_I^*]}{M_0^{*-\sigma}} \right)^{-1} \\ (1 - \tau_0^*) &= \left(\frac{\left[\sum_{i=1}^I \delta^{i-1} \left(M_i \frac{Q_{i-1}^*}{Q_0^*} \right)^{1-\sigma} + \delta^I D^* \left(\frac{Q_I^*}{Q_0^*} \right)^{1-\sigma} \right]}{M_1^{*-\sigma}} \right)^{-1}, \end{aligned}$$

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

B.4 Iteration

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

$$c_0^* = Rb \frac{(R^{I+1} - 1)}{R^{I+1}} \frac{1}{\sum_{i=0}^I \left(\frac{(R\delta)^{\frac{1}{\sigma}}}{R} \right)^i},$$

for all $i \in \{0, \dots, I-1\}$, $c_{i+1}^* = c_i^* (R\delta)^{\frac{1}{\sigma}}$,

$$b_0^* = Rb - c_0^*,$$

for all $i \in \{0, \dots, I-1\}$, $b_{i+1}^* = Rb_i^* - c_{i+1}^*$.

2. Start with a guess for the efficient tax system $\tau = (\tau_0, \dots, \tau_I)$, 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:

$$M_0 = \frac{c_0^*}{Rb(1 - \tau_I) + T_I + G_I} = \frac{c_0^*}{Rb + G_I},$$

For all $i \in \{0, 1, \dots, I-1\}$, $M_{i+1} = \frac{c_{i+1}^*}{Rb_i^*(1 - \tau_i) + T_i + G_i} = \frac{c_{i+1}^*}{Rb_i^* + G_i}$

where

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

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

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

4. Compute D and D_1, \dots, D_I according to (19) and (18) evaluated at the tax guess.
5. Now use the linear policies computed in step 3 and the value function constants computed in step 4 to compute taxes according to the system of equations describing optimal taxes (20).
6. If the taxes you compute in step 5 is the same as the taxes you started the last iteration, stop. If not, use the taxes you computed in step 5 as the new guess and continue iteration.

C Assessing bounds for the parameter β_1

In this section, we explain how we refer to the work of Green, Myerson, and Ostraszewski (1999) to determine a set of reasonable parameter values for the level of self-control problems of agents in period one of our model, β_1 . Green, Myerson and Ostraszewski use experimental data collected from a group of young adults (they also separately study children and older adults) to estimate their temporal discounting function. The group of young adults has a mean age of 20, which corresponds to the age of the agents in period one in our model.

They find that the following hyperbola-like function provides the best description for young adults' discounting functions:

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

where D is the length of delay to a future reward (measured in years) and k and s are parameters which they estimate to be $k = 0.618$ and $s = 0.368$ for the young adult group. This function was originally proposed by Green, Fry, and Myerson (1994).

Our aim is to find the best approximation to this estimated hyperbolic discount function among the set of quasi-hyperbolic discount functions that are parameterized by two parameters, δ and β . As we do all throughout the paper, we follow Laibson, Repetto, and Tobacman (2007) and set $\delta = 0.96$, and choose β that best matches the hyperbola estimated by Green, Myerson and Ostraszewski for young adults. To do so, we first plot the discount factor as a function of years of delay implied by the hyperbolic discount function given above. This is given by the thick solid black line in Figure 11 below. Then, we generate yearly discounting by using $\delta = 0.96$ and $\beta = 0.2$ in the quasi-hyperbolic discounting function. This is the dashed blue line in Figure 11. Then, we increase β to 0.25, and find out that this quasi-hyperbolic approximation of the thick black line, which is given by the green dashed line, dominates the one with $\beta = 0.2$ since it lies closer to the black line for every value on the x-axis. Thus, 0.2 cannot be a plausible value for β of young adults. Then, we keep generating new approximations by increasing β to 0.3, 0.4, 0.5, 0.6, 0.7, 0.75 and 0.8. In the figure, these different approximations are shown by dashed lines of various colors, which shift upwards with every time β increases. With every increase in β up to and including 0.7, the approximation gets closer to the black line for the vertical part of the black line but gets further away regarding the horizontal part. Thus, we cannot dismiss any of these approximations the way we dismissed the approximation generated by $\beta = 0.2$. From $\beta = 0.7$ to 0.75 and to 0.8, the approximation gets worse regarding both the vertical and horizontal segments of the black line. This means $\beta = 0.75$ approximation and approximations with β values higher than 0.75 are clearly dominated by $\beta = 0.7$. Therefore, we conclude the range of β that is reasonable for the young agents is between 0.25 and 0.7.

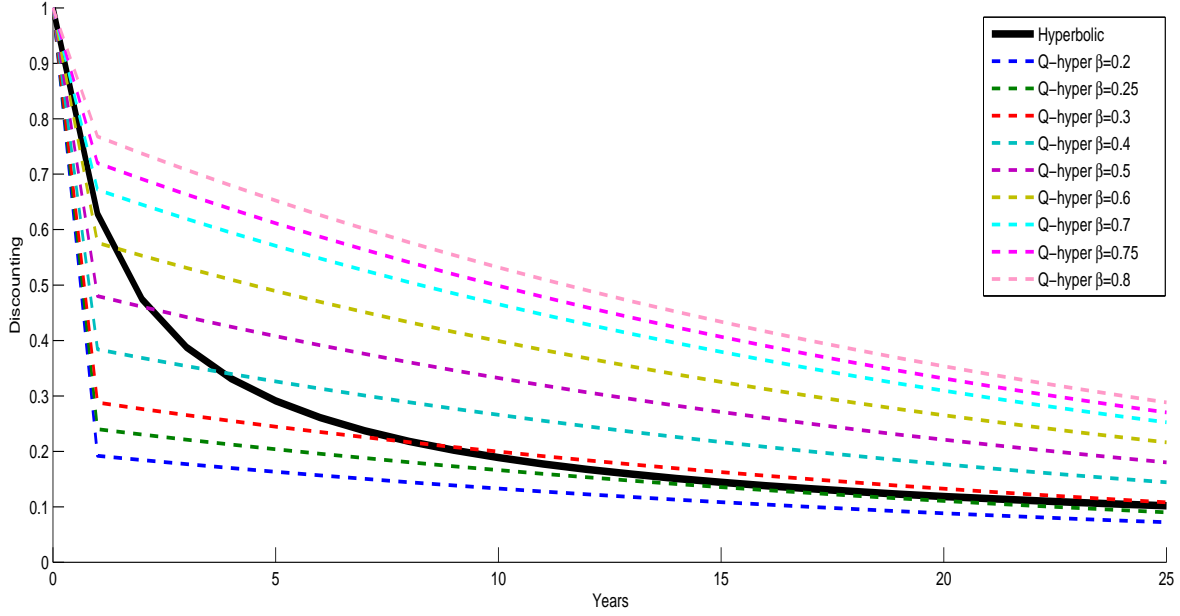


Figure 11: Assessing the upper and lower bounds for β_1 .

D Introducing an Illiquid Asset

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

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

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:

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

Case 1. Government sets taxes such that

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

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

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

efficiency requires

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

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

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

where $\frac{\partial y_1(b_1, d_1)}{\partial b_1} = R(1 - \tau_1)$ (For ease of exposition, assume the policies are differentiable).³¹ Therefore,

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

which implies:

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

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

Case 2. Government sets taxes such that

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

³¹It is well-known that in general we cannot guarantee even the continuity of the policy functions (e.g., see Krusell and Smith (2003), and Harris and Laibson (2001)).

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

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

still holds, efficiency still requires

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

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

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

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

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

which implies:

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

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