Optimal Bequest Taxation and Capital Subsidies Over the Life-Cycle

Nicola Pavoni and Hakki Yazici

November 20, 2011

ABSTRACT:

We study optimal bequest and capital taxation in an economy where agents face self-control problems. In contrast to the previous literature and in line with evidence suggested by personality psychology and experimental studies, we assume that the severity of the self-control problem changes over the life-cycle. We also allow for age-dependent partial sophistication.

We restrict attention to CIES utility functions and derive explicit formulas which allow us to compute optimal taxes given the evolution of self-control problem over the life-cycle. Optimal taxes are very sensitive to the life-cycle pattern of self-control. In particular, if, as suggested by the available empirical evidence, agents' ability to self-control increases concavely with age, then capital should be subsidized and the subsidy should decrease with age.

We furthermore show that (i) if the utility function is further restricted to logarithmic function, then optimal taxes are independent of the life-cycle pattern of partial sophistication; (ii) when agents are assumed to be fully sophisticated and a steady-state condition holds, then optimal taxes are independent of CIES coefficient.

Finally, we use a battery of numerical simulations for several empirically plausible levels of the CIES coefficient and find that - as long as people's level of sophistication does not change abruptly from one period to another – optimal taxes we compute are robust to different patterns of partial sophistication over the life-span.

1 Introduction

Economists traditionally assume that people discount streams of utility over time exponentially. An important qualitative feature of exponential discounting is that it implies time-consistent intertemporal preferences: A person feels the same about a given intertemporal tradeoff no matter 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

[1] See DellaVigna (2009) for a survey of field studies and Frederick et al. (2002) for a survey of experimental studies.

normative implications.[2]

In this paper, we study optimal bequest and capital taxation over the life cycle in the presence of self-control problems. A common modelling assumption in the literature is that the degree of self-control problem is constant over time. Even though this assumption might be a good approximation of reality for analyzing many questions, we believe it is not suitable for life cycle analysis because a significant body of empirical studies points to the opposite: like many other personality traits, people's ability to self-control changes as they age. The first set of evidence for changing level of self-control over the life-span comes from personality psychology. As Ameriks et al. (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 hence the degree of self-control, changes with age.[4] Indeed, in a survey article on personality development in adulthood, Roberts et al. (2003) conclude "it appears that the increase in conscientiousness is one of the most robust patterns in personality development, especially in young adulthood." There is a second set of more direct evidence in favor of changing self-control: research on intertemporal discounting over the life-span has shown that short term discount rates fall with age predicting a a life cycle developmental trend toward increased self control.[5] All this evidence suggests that, in order to study bequest and capital taxation over the life cycle, one should extend the traditional models of self-control to allow for varying degrees of self-control problem over the life cycle. This is exactly what the current paper does.

In our model, in each period agents make consumption savings decisions facing age-dependent degrees of self control problems and sophistication, which is defined as people's awareness of their self-control problems. Later in life, people become altruistic parents, make bequests, and die. We model preferences that exhibit self-control problems through the hyperbolic discounting model of Laibson (1997), which builds on the seminal works of Strotz (1956) and Phelps and Pollak (1968). We extend the Laibson (1997) model in two ways that are crucial to our analysis. First, we allow for the degree of self-control problem to change over time. Second, we introduce partial sophistication to that model which essentially amounts to allowing for different degrees of self awareness about the existence of the self-control problems.[6] We allow for the degree of partial sophistication to change over time. In this environment, we define the first-best allocation as the allocation that would arise in the absence of self-control problems. The main exercise in this paper is to examine the optimal tax policy that implements the first-best allocation.

The results are quite striking. First, we show that optimal bequest taxes are always strictly positive. This is a very

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

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

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

[5] Green et al. (1994);(1999), and Read and Read (2004).

[6] We are not the first ones to model partial sophistication, O'Donoghue and Rabin (1999) is. However, the way we introduce partial sophistication is significantly different from theirs. We justify our way of modeling partial sophistication on the grounds of tractability and the fact that the two models deliver very similar predictions.

general result as it does not depend on how self-control problem and sophistication evolve over the life cycle. Second, we examine optimal capital taxes along the life cycle and show that capital taxes should be age-dependent. Regarding the signs of capital taxes at different ages and the monotonicity of capital taxes with respect to age, we prove an ambiguity result: 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. We believe this is an important message since it shows that researchers who take self-control problems seriously should also take the evolution of self-control problems over the life-cycle seriously before making policy suggestions. This result is also interesting because it invalidates the basic presumption in the literature that self-control problems always imply subsidies.[7] The previous literature missed this result because of their assumption of constant self-control problems which always implies subsidies.

Next, we restrict attention to constant relative risk aversion (CRRA) instantaneous utility functions and provide sharper results about the properties of age-dependent capital taxes. First, for logarithmic utility, we derive closed form formulas for optimal taxes. Then, we prove that, when utility is logarithmic, optimal taxes are independent of how sophistication changes over the life cycle. Second, we show that if a steady-state condition is satisfied and agents are fully sophisticated, then optimal taxes are independent of CRRA coefficient σ. These results make the tax formulas computed for the logaritmic case quite general: they are valid under any pattern of partial sophistication when utility is logarithmic and under any σ when agents are fully sophisticated. Using these formulas, we prove that if, as strongly suggested by personality psychologists, the degree of self-control increases with age, then capital should be subsidized in all periods. Finally, 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.

Finally, we know from O'Donoghue and Rabin (1999) that allowing for even constant level of partial naivete can change people's behavior, and hence, optimal policy significantly.[8] We analyze how changing naivete over the life-span alters our optimal taxation results. It is evident then 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. Unfortunately, when $\sigma \neq 1$ and agents are allowed to be partially sophisticated, we do not get closed for solutions for optimal taxes. Therefore, we resort to numerical analysis. We derive two conclusions from our numerical experiments. First, as long as the level of sophistication is not changing abruptly from one period to another, the pattern of optimal capital subsidies over the lifecycle is surprisingly robust to how sophistication changes with age. Second, this result is independent of σ, which implies that the pattern of optimal capital subsidies is robust to σ, as long as the level of sophistication does not jump from one period to another.

Krusell et al. (2002) and Krusell et al. (2010) are the two most closely related papers to the current study. The first one, Krusell et al. (2002), considers infinitely living agents facing self control problems in the form of quasi-hyperbolic discounting a la Laibson. They restrict attention to logarithmic utility and show that a constant subsidy to investment

[7] O'Donoghue and Rabin (1999) is an exception where it says if the agent is sophisticated then he may oversave. However, even in that paper it says that "naifs will undersave in essentially any savings model" and hence should be subsidized. We show that even naifs may oversave and hence may need to be taxed.

[8] See Ariely and Wertenbroch (2001) for behavioral evidence on partial sophistication.

(similar to our capital subsidy) implements the commitment allocation. Krusell et al. (2010) analyses optimal taxation in an economy where agents live finitely many periods and have temptation and self control problems a la GP. They first prove that under CRRA preferences, as the parameter that controls temptation goes to infinity, the optimal policy prescriptions of the quasi-hyperbolic model and the temptation model become identical. Then, they show that for the logarithmic utility special case this equivalence result holds for at any temptation level and, go on to compute optimal savings taxes under logarithmic utility. They show that savings should be subsidized and that this subsidy should be increasing with time due to finite life time effect.[9] Our paper differs from these two papers in significant ways. First and foremost, we allow for changing level of self-control problems over the lifecycle while both papers assume the level of self control problem to be constant over time. By assuming empirically plausible patterns of self control problems over the lifecycle, we show that capital subsidies should actually be decreasing with age. Second, we allow for agents to be partially aware of their future self control problems (partial sophistication) as opposed to assuming people at all ages predict their future self control level perfectly which is the assumption these papers make. This allows us to study the affects of sophistication on capital subsidies. Finally, we assume an intergenerational structure which allows us to make policy predictions about bequest taxation.

2 Model

The economy is populated by a unit measure of dynasties who live for a countable infinity of periods, $t = 1, 2, ...$ Each agent within a dynasty is active for $I+1$ periods: I periods of young adulthood, which lasts until the young adult becomes a parent, and parenthood. In any period, there is only one generation alive; so this is a non-overlapping generations model. During young adulthood, agents make consumption-saving decisions facing different degrees of self-control problems at different ages. Later in life people become parents and make bequests. Parents are altruistic, do not face self-control problems, and are sophisticated in the sense that they anticipate their children's self-control problems.[10] We use hyperbolic discounting formalized by Laibson (1997) to model self-control problems as follows.

A parent's preference over dynastic consumption stream is given by:

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

where u is instantaneous utility function with usual properties, $u', -u'' > 0$ and δ is the discount factor. Period 0 represents the period of parenthood after which the parent dies and is replaced by his offspring who is now in his 1^{st} period of young adulthood. Period i represents the i^{th} period of the offspring. Throughout the paper, we call the agent in his i^{th} period of young adulthood self i. c_0 is current parent's consumption level, c_i self i's consumption level, and c_0' is the parental consumption level of the offspring. The offspring has different preferences at different periods of his life:

[9] In the infinite horizon version of their model, the subsidies would be constant as they show in their working paper Krusell et al. (2009).

[10] As we explain later, the fact that parents have zero self-control problem and that they perfectly anticipate the future behavior of their followers is not crucial to our results.

$$
\begin{aligned}
&u(c_1) + \beta_1 \delta \left[u(c_2) + \delta u(c_3) + \delta^2 u(c_0') + ... \right], \\
&u(c_2) + \beta_2 \delta \left[u(c_3) + \delta u(c_0') + ... \right], \\
&. \\
&. \\
&u(c_I) + \beta_I \delta \left[u(c_0') + ... \right].
\end{aligned}
$$

Here the first equation is young adult'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 fully rational; there is no self-control problem. Throughout the paper we will assume that $\beta_i < 1$, meaning young adults 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, as people age, the severity of the self-control problem they face changes. Therefore, we extend the hyperbolic discounting model by allowing for the severity of self-control problems, β_i, to depend on i.

The other 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. The self-control literature has mostly analyzed two extreme cases regarding this dimension: one in which people are fully naive and another in which they are fully sophisticated. In the naive model, agents believe their followers do not have any self-control problems whereas in the sophisticated model agents are aware of the full extent of the future self-control problems. We allow for partial sophistication which essentially amounts to allowing for different degrees of self awareness about the existence of self-control problems.[11] We extend the hyperbolic discounting model by introducing partial sophistication and allowing the level of sophistication to change over time. We explain the way we model partial sophistication formally when we display the parent's problem.

Production takes place at the aggregate level according to the function $f(k) = Rk$, where k is aggregate capital and f is a neoclassical concave production function.[12] There is a credit market in which agents can trade one-period risk-free bonds. Since all agents are identical at any time, so are their bond holdings. Let b_i be the amount of bond holdings of the agent at age i. Aggregating these bond holdings gives us the aggregate level of capital in each period, so the capital market clearing condition is $k_i = b_i$.

2.1 First-Best Allocation

The first best allocation is the allocation that arises if no one in the economy had self-control problems. It is given by the solution to a fictitious social planner's consumption-saving problem where the planner has discounting with δ. The

[11] We are not the first ones to model partial sophistication, O'Donoghue and Rabin (1999) is. However, the way we introduce partial sophistication is significantly different from theirs. We justify our way of modeling partial sophistication on the grounds of tractability and the fact that the two models deliver very similar predictions.

[12] We could also have $f(k, l)$, where l is aggregate labor and is inelastically supplied. This does not change the current results but allows us to study price externality implicit in self-control environment with endogenous prices, as shown in Krusell et al. (2002).

following Euler Equations characterize the first-best allocation, which we denote with superscript star throughout the paper:

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

Also, observe that steady-state requires $\delta R = 1$.

2.2 Implementing First-Best via Bequest and Capital Taxes

Since young agents in this economy have self-control problems, laissez-faire market equilibrium cannot attain first-best allocation. In this subsection, we define market equilibrium with taxes. We call a tax system *optimal* if it implements the first-best allocation in the market environment.

2.2.1 Parent's Problem with Taxes

The following is the parent's problem under taxes. τ_i is the savings (capital) tax agent $i = 0, 1, .., I$ pays. Tax proceeds are rebated in a lump-sum manner in every period, $T_i = R\tau_i b_i$ so that government balances its budget period by period.

$$\begin{aligned}
W(b;\boldsymbol{\tau}) &= \max_{b_0,b_1,..,b_I} \; u\left(R(1-\tau_I)b + T_I - b_0\right) + \delta\left[\sum_{i=0}^{I-1}\delta^i u\left(R(1-\tau_i)b_i + T_i - b_{i+1}\right) + \delta^I V(b_I;\boldsymbol{\tau})\right] && (1)\\
& s.t. \\
(b_1,..,b_I) &\in \arg\max_{\hat{b}_1,..,\hat{b}_I} \; u\left(R(1-\tau_0)b_0 + T_0 - \hat{b}_1\right) && \text{(self-1)}\\
&+\delta\beta_1\left[\begin{array}{c}\pi_1\left\{\sum_{i=1}^{I-1}\delta^{i-1}u\left(R(1-\tau_i)\hat{b}_i + T_i - \hat{b}_{i+1}\right) + \delta^{I-1}V\left(\hat{b}_I;\boldsymbol{\tau}\right)\right\}\\ +(1-\pi_1)W_1\left(\hat{b}_1;\boldsymbol{\tau}\right)\end{array}\right] \\
& s.t \\
\left(\hat{b}_2,..,\hat{b}_I\right) &\in \max_{\left(\bar{b}_2,..,\bar{b}_I\right)} \; u\left(R(1-\tau_1)b_1 + T_1 - \bar{b}_2\right) && \text{(self-2)}\\
&+\delta\beta_2\left[\begin{array}{c}\pi_2\left\{\sum_{i=2}^{I-1}\delta^{i-2}u\left(R(1-\tau_i)\bar{b}_i + T_i - \bar{b}_{i+1}\right) + \delta^{I-2}V\left(\bar{b}_I;\boldsymbol{\tau}\right)\right\}\\ +(1-\pi_2)W_2\left(\bar{b}_2;\boldsymbol{\tau}\right)\end{array}\right] \\
& s.t \\
& ... \\
& s.t. \\
\left(b'_{I-1},b'_I\right) &\in \max_{\tilde{b}_{I-1},\tilde{b}_I} \; u\left(R(1-\tau_{I-1})b_{I-2} + T_{I-2} - \tilde{b}_{I-1}\right) \\
&+\delta\beta_{I-1}\left[\pi_{I-1}\left\{u\left(R(1-\tau_{I-1})\tilde{b}_{I-1} + T_{I-1} - \tilde{b}_I\right) + \delta V\left(\tilde{b}_I;\boldsymbol{\tau}\right)\right\} + (1-\pi_{I-1})W_{I-1}\left(\tilde{b}_{I-1};\boldsymbol{\tau}\right)\right] \\
& s.t \\
\tilde{b}_I &\in \max_{b''_I} \; u\left(R(1-\tau_{I-1})b_{I-1} + T_{I-1} - b''_I\right) + \delta\beta_I\left[\pi_I V\left(b''_I;\boldsymbol{\tau}\right) + (1-\pi_I)W_I\left(b''_I;\boldsymbol{\tau}\right)\right] && \text{(self-I)}
\end{aligned}$$

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

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

with

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

Here $W(b; \boldsymbol{\tau})$ represents the value of a parent's problem who saved b units in his last period before parenthood and faces the tax system $\boldsymbol{\tau}$. The parent chooses his bequest b_0 and life cycle savings for his offspring $b_1, ..., b_I$. However, since the parent does not have any direct control over $b_1, ..., b_I$ and his preferences are not aligned with his offspring's (in a given period i, parent's discount factor is δ whereas offspring's is $\beta_i \delta$), parent's choice of $b_1, ..., b_I$ has to be incentive compatible from the perspective of his offspring's selves. The parent is sophisticated in the sense that he foresees this misalignment of preferences, and that is the reason why he takes into account the incentive compatibility conditions we see in his problem.

To understand the nested nature of incentive compatibility conditions better and the way we model partial sophistication, let us analyze the last two incentive constraints. First, the last contraint describes how self I chooses b_I. π_I 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 is going to face offsprings with self-control problems, i.e. $(\beta_1, .., \beta_I) \neq 1$. Note that in reality this probability is one; however, self I is partially naive in the sense that he incorrectly attaches positive probability $(1 - \pi_I)$ to the event that there will never be self control problems in the future, i.e. $(\beta_1, .., \beta_I) = 1$. So, in our environment, π_I represents the level of sophistication of self I. We assume that all the agents, including the parents, correctly guess the level of sophistication of their future selves: so parent knows π_I and b_I is chosen in reality. Second, consider the incentive constraint that displays how self $I-1$ chooses b_{I-1} and b_I. π_{I-1} represents the degree of sophistication of self $I-1$. Note that from the perspective of self $I-1$, self I has a self control problem, and self $I-1$ is partially sophisticated, meaning he knows with probability π_{I-1} that self I is going to have self-control problem when he is choosing b_I. Thanks to our assumption that all agents know their future selves' level of sophistication, with probability π_1, self $I-1$ perceives correctly how self-I is going to choose b_I, which is described in (1). We have just seen that the last incentive constraint enters the parent's problem in at least two ways: directly as self I's incentive constraint and as a constraint of self $I-1$'s incentive constraint. So, we do not impose the incentive constraint of self I separately on the parent's problem. Using the exact same logic, one can show that self $I-1$'s incentive constraint is also a constraint in self $I-2$'s incentive constraint, and self $I-2$'s incentive constraint is also a constraint of self $I-3$'s, and so on. Thus, incentive constraint of self i enters parent's problem in i different places: directly and as constraints of previous i-1 selves' incentive constraints. As a result, the incentive constraint of self 1 includes the incentive constraints of all future selves and hence is sufficient to summarize future behavior, $b_1, ..., b_I$, given b_0.

Using the first order approach, we can replace agents incentive constraints with associated first order conditions:

$$
\begin{aligned}
W(b;\boldsymbol{\tau}) &= \max_{b_0,b_1,..,b_I} \; u\left(R\left(1-\tau_I\right)b+T_I-b_0\right)+\delta\left[\sum_{i=0}^{I-1}\delta^i u\left(R\left(1-\tau_i\right)b_i+T_i-b_{i+1}\right)+\delta^I V\left(b_I;\boldsymbol{\tau}\right)\right] && (2)\\
& s.t.\ for\ all\ i\in\{0,...,I-2\} && (3)\\
u'(R(1-\tau_i)b_i+T_i-b_{i+1}) &= \delta\beta_{i+1}\left[\begin{array}{c}\pi_{i+1}\left\{\begin{array}{c}\sum_{j=i+1}^{I-1}\delta^{j-(i+1)}u'(R(1-\tau_j)b_j+T_j-b_{j+1})\left(R(1-\tau_j)\frac{\partial b_j}{\partial b_{i+1}}-\frac{\partial b_{j+1}}{\partial b_{i+1}}\right)\\ +\delta^{I-(i+1)}V'(b_I;\boldsymbol{\tau})\frac{\partial b_I}{\partial b_{i+1}}\end{array}\right\}\\ +(1-\pi_{i+1})W'_{i+1}(b_{i+1};\boldsymbol{\tau})\end{array}\right] && (4)\\
& and\ s.t. && (5)\\
u'\left(R\left(1-\tau_{I-1}\right)b_{I-1}+T_{I-1}-b_I\right) &= \delta\beta_I\left[\pi_I V'\left(b_I;\boldsymbol{\tau}\right)+\left(1-\pi_I\right)W'_I\left(b_I;\boldsymbol{\tau}\right)\right]
\end{aligned}
$$

We restrict attention to equilibrium that is the limit of the unique finite-horizon equilibrium. This restriction gives a unique equilibrium with policy functions that are linear in net present value of current wealth:

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

where G_i is the net present value of future lump-sum taxes, $\Gamma_{i-1}(b_{i-1})$ is the net present value of wealth available to agent at age i, and M_i is the fraction consumed out of that wealth. It is relatively simple algebra to show that:

$$
\Gamma_i(b_i)=R(1-\tau_i)(1-M_i)\Gamma_{i-1}(b_{i-1})
$$

Using linearity of the policy functions, we can rewrite the problem as:

$$
\begin{aligned}
W(b;\boldsymbol{\tau}) &= \max_{M_0,..,M_I} \; u\left(M_0\Gamma_I(b)\right)+\delta\left[\sum_{i=1}^{I-1}\delta^i u\left(Q_{i-1}M_i\Gamma_I(b)\right)+\delta^I V\left((1-M_I)Q_{I-1}\Gamma_I(b);\boldsymbol{\tau}\right)\right] && (6)\\
& s.t.\ for\ all\ i\in\{1,...,I-1\}\\
\left(M_iQ_{i-1}\Gamma_I(b)\right)^{-\sigma} &= \delta\beta_i\left[\begin{array}{c}\pi_iR(1-\tau_i)\left\{\sum_{j=i+1}^{I}\delta^{j-(i+1)}\left(M_jQ_{j-1}\Gamma_I(b)\right)^{-\sigma}M_j\frac{Q_{j-1}}{Q_i}+\delta^{I-i}V'(b_I;\boldsymbol{\tau})(1-M_I)\frac{Q_{I-1}}{Q_i}\right\}\\ +\left(1-\pi_i\right)W'_i\left(b_i;\boldsymbol{\tau}\right)\end{array}\right] && (7)\\
& and\ s.t. && (8)\\
\left(M_IQ_{I-1}\Gamma_I(b)\right)^{-\sigma} &= \delta\beta_I\left[\pi_I V'\left(b_I;\boldsymbol{\tau}\right)+\left(1-\pi_I\right)W'_I\left(b_I;\boldsymbol{\tau}\right)\right] && \text{(self-3)}
\end{aligned}
$$

where

$$
Q_{i-1}:=R^i\Pi_{s=0}^{i-1}\left(1-M_s\right)\left(1-\tau_s\right).
$$

3 A Simple Example

In this section, we consider a simple example. The simplest environment to analyze the main mechanism has to consist of 3 periods since we need an agent, self 1, choosing a savings level, b_1, taking into account the action of a future self, self 2 choosing b_2, and future government policies, τ_2.

The problem of self 1 is:

$$\begin{aligned} & \max_{b_1,b_2,\hat{b}_2} u(w-b_1)+\beta_1\delta\{\pi_1\left[u(R(1-\tau_1)b_1+T_1-b_2)+\delta u(R(1-\tau_2)b_2+T_2)\right] \\ & +(1-\pi_1)\left[u(R(1-\tau_1)b_1+T_1-\hat{b}_2)+\delta u(R(1-\tau_2)\hat{b}_2+T_2)\right]\} \\ & s.t. \\ b_2 \in \; & \arg\max_{\tilde{b}_2} u(R(1-\tau_1)b_1+T_1-\tilde{b}_2)+\beta_2\delta u(R(1-\tau_2)\tilde{b}_2+T_2) \end{aligned}$$

where b_1, b_2 are self 1 and self 2's savings and $\hat{b}_2$ represents what self 1 naively believes self 2 will choose.

The FOC of the partially sophisticated self 1 is:

$$\begin{aligned} u'(c_1) &= \pi_1\delta\left[u'(c_2)[R(1-\tau_1)-b_2']+\delta R(1-\tau_2)b_2'u'(c_3)\right] \\ &\quad +(1-\pi_1)\delta\left[u'(\hat{c}_2)[R(1-\tau_1)-\hat{b}_2']+\delta R(1-\tau_2)\hat{b}_2'u'(\hat{c}_3)\right] \\ &= \beta_1\delta R(1-\tau_1)u'(c_2)\left\{\pi_1\left[1+b_2'(b_1)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1)}\right]+(1-\pi_1)\frac{u'(\hat{c}_2)}{u'(c_2)}\right\}, \end{aligned}$$

where $\hat{c}_2$ represents self 1 naive belief about self 2's consumption choice. This means that optimal tax is given by:

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

The tax formula above consists of two components. The first part, $\frac{1}{\beta_1}$, is easier to understand. Because of his *current* self control problem, self 1 discounts tomorrow by an extra β_1 and hence wants to undersave relative to the planner. By multiplying the after tax return with $\frac{1}{\beta_1}$, we can exactly offset the extra discounting, thereby getting rid of this undersaving motive of the agent. Let us call this first part of the tax formula the *current component*. Obviously, current component is not affected by sophistication level at all. Clearly, the current component of tax is always negative, i.e. it always calls for a subsidy. This is not the end of the story however. Self 1's choice of current savings is also affected by the actions of future selves and future government policies. Therefore, even if we correct for his undersaving through the current component of the tax, he still deviates from first-best saving level in order to compensate for his future selves' suboptimal actions (due to future self control problems) and/or in response to future policies. The second part of the tax formula $\left\{\pi_1\left[1+b_2'(b_1)\frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1^*)}\right]+(1-\pi_1)\frac{u'(\hat{c}_2)}{u'(c_2)}\right\}^{-1}$ is there to correct deviations in current savings caused by future actions and policies. We call this part the *future component* of the tax formula. This is where the level of sophistication matters. As we will see, the future component of taxes is always positive. The sign of the optimal capital tax can be positive or negative depending on which component dominates.

In the rest of this section we analyze the future component closely. Therefore, to isolate the future component, we assume $\beta_1 = 1$.

Sophisticated Future Component:

Let's first analyze the future component when the agent is fully sophisticated in period one, so suppose $\pi_1 = 1$. We call the tax on sophisticated period one agent τ_1^S. FOC of self 2 reads:

$$u'(c_2) = \beta_2 \delta R(1-\tau_2)u'(c_3). \tag{9}$$

This implies that in order to implement first best period two saving, b_2^*, the planner has to subsidize the return to period two saving by $(1-\tau_2^*) = \frac{1}{\beta_2}$. Let $b_2(b_1)$ be the solution to self 2's problem.

Self1's FOC reads:

$$u'(c_1) = \delta \left[u'(c_2)R(1-\tau_1) + b_2'(b_1) \left\{ -u'(c_2) + \delta R(1-\tau_2)u'(c_3) \right\} \right]. \tag{10}$$

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

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

which implies

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

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

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

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

Naive Future Component:

Now lets analyze the future component when self 1 is fully naive, so suppose $\pi_1 = 0$. We call the tax on the naive self 1 agent τ_1^N. Let $b_2(b_1)$ and $\hat{b}_2(b_1)$ be self 2's actual choice and naive self 1's expectation of self 2's choice, respectively. The FOC for self 2 is:

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

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

On the other hand, self 1 incorrectly believes that self 2 chooses his savings according to:

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

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

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

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

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

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

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

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

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

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

3.1 Equivalence in the Log Case

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

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

Sophisticated future component:

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

Naive future component:

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

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

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

Since the current component of the tax is only related to agent's current level of self-control problem and is given by $\frac{1}{\beta_i}$, under logarithmic utility optimal taxes are independent of the degree of sophistication. We begin next section by generalizing this result to our full blown environment.

4 Optimal Taxes

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

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

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

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

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

Proof. Relegated to Appendix A. ■

The proposition below shows that if $\delta R=1$, meaning we are at a 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 *If $\delta R=1$ and $\pi_i=1$ for all i, then optimal taxes are independent of CRRA coefficient σ.*

Proof. Relegated to Appendix A. ■

The two propositions above imply that the tax formulas of proposition 1 are quite general.

4.1 Lessons for Bequest and Capital Taxation

This proposition implies several important and general lessons for taxes which are summarized in a series of corollaries below.

Corollary 3 ***(Sign of the Bequest Tax)*** *Optimal bequests taxes are always positive, independent of how the degree of self control problem and sophistication evolves over the life cycle.*

The first lesson to be taken is very general as it does not depend on the utility function specification as well: bequests should always be taxed. To see the intuition behind this result, remember that in the previous subsection we broke down the expression for taxes into two components: current and future. Regarding the future component of bequest tax, we need to remember the discussion on sophisticated future component since parents are assumed to be sophisticated. The idea here is that as long as the offspring has a self-control problem in the first period of his young adulthood, from parent's perspective he is undersaving. To compensate for this, the parent oversaves which implies he should be taxed. The current component is zero since we assume parents have no self-control problems, which means the future component dominates and we get positive tax on bequests.[13] It is important to realize that sophistication of parents is not needed for the positive bequest tax result as future component of savings tax is always positive independent of sophistication, but the intuition behind the tax result would be different if parents are (partially) naive. Furthermore, the assumption that parents have zero self-control problems is probably also not needed; as the logaritmic utility case suggests, we only need self-control to be low at old age.

Corollary 4 ***(Age-dependence)*** *Optimal capital taxes are age-dependent.*

The second lesson to be learnt from logarithmic utility case is that in general optimal capital taxes should depend on people's age.[14] The reason for the necessity of this dependence is the changing the degree of self-control problem over age, which is a well established fact in personality psychology literature as discussed in the introduction.

Corollary 5 ***(Sign of the Capital Taxes)*** *(1) Optimal capital taxes might be positive or negative depending on how β_i change with i.*

(2)With log utility, optimal capital taxes are always negative in the last period before parenthood, $\tau_I < 0$. If $\beta_{i+1} \geq \beta_i$, for all i, then optimal capital tax is negative for all ages:

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

The general lesson to be taken about the sign of the age-dependent capital taxes is simply that optimal capital taxes might be positive or negative depending on the evolution of the severity of the self-control problem over the life cycle. This is an important message since it shows that researchers who take self-control problems seriously should also take the evolution of self-control problems over the life-cycle seriously before making policy suggestions. This is quite contrary to

[13] Parents leave bequests and die. Therefore, bequest decision depends on parent's relative preference of his consumption over his offspring's. Therefore, the assumption that parents have no self-control problems amounts to assuming parent's preference towards his offspring's consumption does not change over time. Even though we believe this is a plausible assumption, it is possible to show that we do not need it. Suppose parents face self-control problem, the level of which is denoted by β_0. Then, the bequest tax will be given by

$$1 - \tau_0 = \frac{1}{\beta_0}(1 - \delta + \beta_1\delta).$$

A sufficient condition for the above expression to be strictly less than 1 (and hence for a positive tax) is $1 - \delta < \beta_0 - \beta_1$. Assuming a period in the model is a year and yearly discount function is around 0.95, this roughly means, $\beta_0 - \beta_1 > 0.05$ is sufficient to get bequest tax. As a justification for this assumption, the β estimates Green et al. (1999) and Read and Read (2004) find for their youngest and oldest groups show that $\beta_0 - \beta_1$ is well above 0.2.

[14] Krusell et al. (2010) also states the need for age-dependence of savings taxes but in their case this is due to a finite life time effect and washes away if one analyzes the infinite version of their economy.

the presumption in the literature that self-control problems always imply subsidies.[15] The reason why previous literature did not see this result is because they assumed constant self-control problems which directly implies subsidies.

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

Corollary 6 ***(Monotonicity of Capital Taxes)*** *(1) Optimal capital taxes might be increasing or decreasing depending on how β_i change with i.*

(2) If $0 \leq \beta_{i+1} - \beta_i \leq \beta_i - \beta_{i-1}$ (concavity) for all i, then optimal capital subsidies decrease with age.

The general lesson about the monotonocity of capital taxes again points to the importance of the evolution of the severity of self control problem over the life cycle: without the knowledge of how β_i changes with i, policy prescriptions would be misguided. With logarithmic utility or CRRA under full sophistication, there is again a sharper message: independent of the level of sophistication, capital subsidies should be decreasing with age if people's self control level is increasing concavely. There is some evidence for this in the personal psychology literature. This result is contrary to Krusell et al. (2010) which concludes that in any finite economy, taxes should be increasing with age.

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

4.2 Quantitative Analysis

In this subsection, we numerically analyze optimal bequest taxes and capital subsidies over the lifecycle assuming that either one of the justifications of the tax formulas of proposition 1 hold: either utility is logarithmic or all the agents in the model are fully sophisticated. In order to a numerical analysis, we have to choose particular values for parameters. 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. We set the true yearly discount factor δ to the inverse of the yearly gross interest rate in the US which is taken to be $R = 1.04$. Due to the proposition above, we know that optimal taxes will not depend on the constant relative risk aversion coefficient σ; so we do not need to specify a value for it.

As it is evident from the optimal tax formulas, self control function $\beta(i)$ is the crucial parameter. Moreover, Figure 1 and Figure 2 show that taxes are in general very sensitive to $\beta(i)$. Therefore, to say something concrete, we need to make several assumptions about $\beta(i)$. We assume that $\beta(i)$ is increasing and concave in i. In words, this means that the degree of self-control problem decreases with age and this decline slows down with age. We have two sets of evidence in favor

[15] O'Donoghue and Rabin (1999) is an exception where it says if the agent is sophisticated then he may oversave. However, even in that paper it says that "naifs will undersave in essentially any savings model" and hence should be subsidized. The log case shows that in our environment even naifs may oversave and hence may need to be taxed.

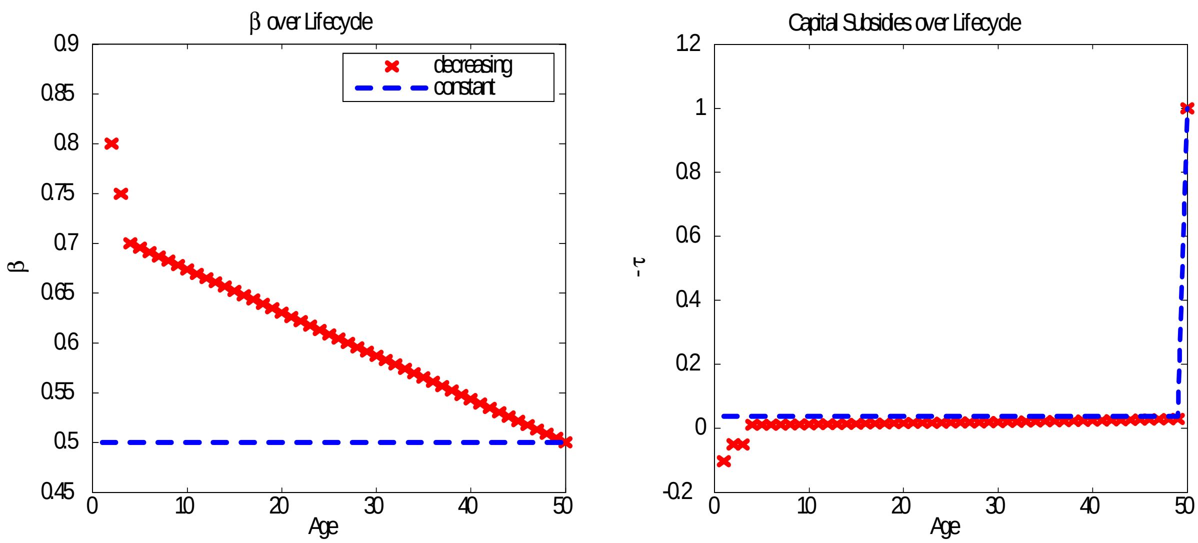

Figure 1: Optimal capital subsidies for decreasing and constant patterns of β over the life-cycle.

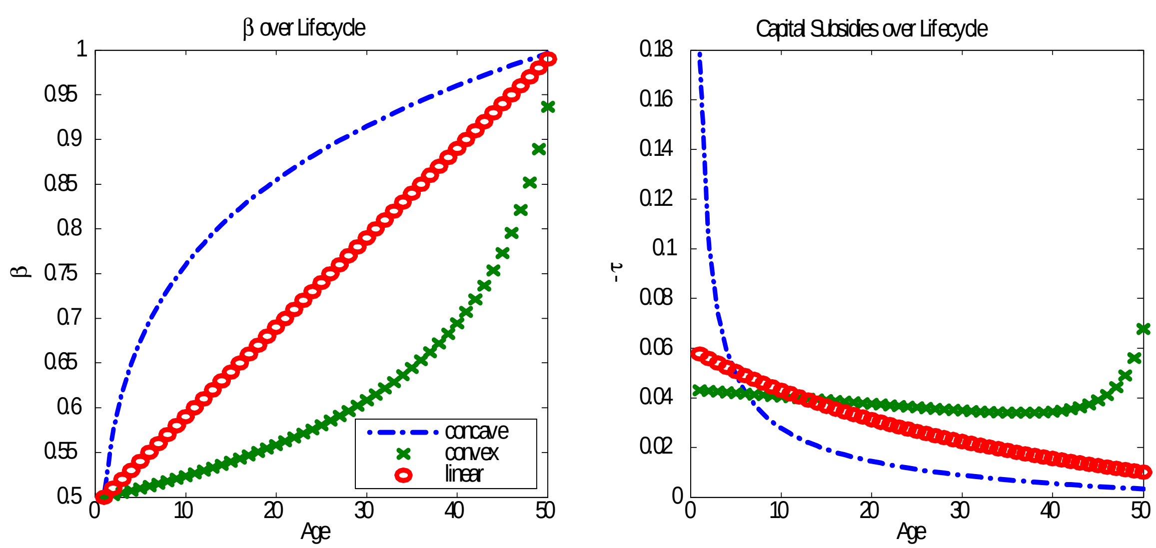

Figure 2: Optimal capital subsidies for concave, linear, and convex increasing patterns of β over the life-cycle.

of these assumptions. First, research on intertemporal discounting over the life-span has shown that short term discount rates fall with age predicting a a life cycle developmental trend toward increased self control. [16] Second, personality psychologists associate self control with conscientiousness, one of the "big five" personality factors,[17] and in the words of Roberts et al. (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. Srivastava et al. (2003) and Roberts et al. (2006) both find that conscientiousness increases concavely over the lifecycle. Srivastava et al. (2003) estimates conscientiousness as a quadratic function of age and finds that the quadratic age term has a negative coefficient "indicating that the rate of increase [in conscientiousness] was greater at younger ages than at older ages."[18] We use a quadratic β function and perform robustness checks by varying the degree of concavity allowing for linearity as well.

We also make assumptions about the level of β at the youngest and oldest age. In most of our calculations, we assume that $\beta(1) = 0.5$. This number is below 0.7 which is what is commonly estimated and used by economists in models with constant β. The reason for such a choice is that our youngest agent is 20 years old which is significantly smaller than the mean age in most of these studies. We check for robustness by setting $\beta(1) = 0.4$ and $\beta(1) = 0.65$.[19] We assume that self control problem vanishes at the end of one's life 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 75 in Read and Read (2004) and 70 in Green et al. (1999). Finally, we assume a quadratic funtional form for the $\beta(i)$ function, following the estimation of conscientiousness as a function of age.

Now we report the results. In all our simulations, capital taxes are negative so they are indeed subsidies and these subsidies are decreasing with age throughout the lifecycle. This is expected since we assumed β is increasing and concave and proposition 1 shows in that case we always have positive, decreasing subsidies. Figure 3 below aims to display the sensitivity of subsidies to the degree of concavity of β function. In the figure, the β function is depicted on the left while the corresponding age-dependent capital subsidies are depicted on the right. The initial level of self-control problem β_1 is set to 0.5. The three curves other than the solid one represent different degrees of concavity within the same family of quadratic functions. The blue dashed line has the highest level of concavity whereas the green straight one has the least, and the red dotted curve is in between. We see that the more concave the function is the higher the initial level of taxes are, the faster the decline with age is, and the lower the value of final level of taxes. In the solid curve in turqoise, β is a 4th root function of age, and this shows that the type of concave function chosen also matters. Figure 4 shows sensitivity

[16] Green et al. (1994) and (1999) and Read and Read (2004). and also ameriks

[17] Ameriks et al. (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." Heckman et al (2008) also states that concsientiousness is conceptually related to self-control problems.

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

[19] 0.65 is the one-year short-term discount rate we computed using Green et al. (1999)'s estimate of hyperbolic discount function for his young adult group which has mean age of 20 years.

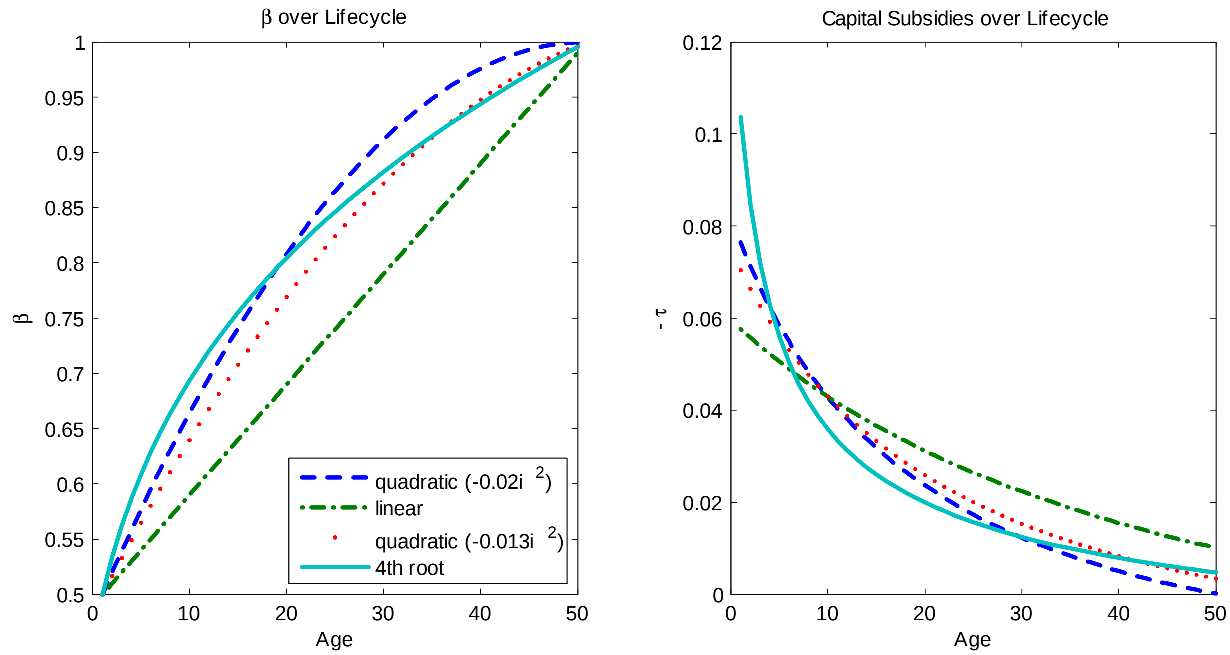

Figure 3: Sensitivity with respect to the curvature of β

with respect to our assumption about the initial level of the self-control problem assuming β is quadratic. We see that the lower is β_1 the higher is the starting value of taxes and also the sharper the decline.

The level of bequest tax is directly linked to the value of β_1 and is summarized in the table below:

β_1	Bequest Tax
0.4	58%
0.5	48%
0.65	34%

5 Effect of Partial Sophistication

In the previous section, we show that: (1) when the constant relative risk aversion coefficient $\sigma = 1$, then the degree of sophistication is immaterial for taxes; (2) under the assumption that all the agents in the economy are fully sophisticated, σ is immaterial for taxes. In these two cases, taxes are given by proposition 1. Then, by making certain plausible assumptions on the β function, we displayed the lifecycle properties of optimal subsidies. It is evident then 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 for solutions for optimal taxes. Therefore, we have to resort to numerical analysis.

In our first set of analysis, we set $\sigma = 2$ and allow for people's level of sophistication to be different from 1 and

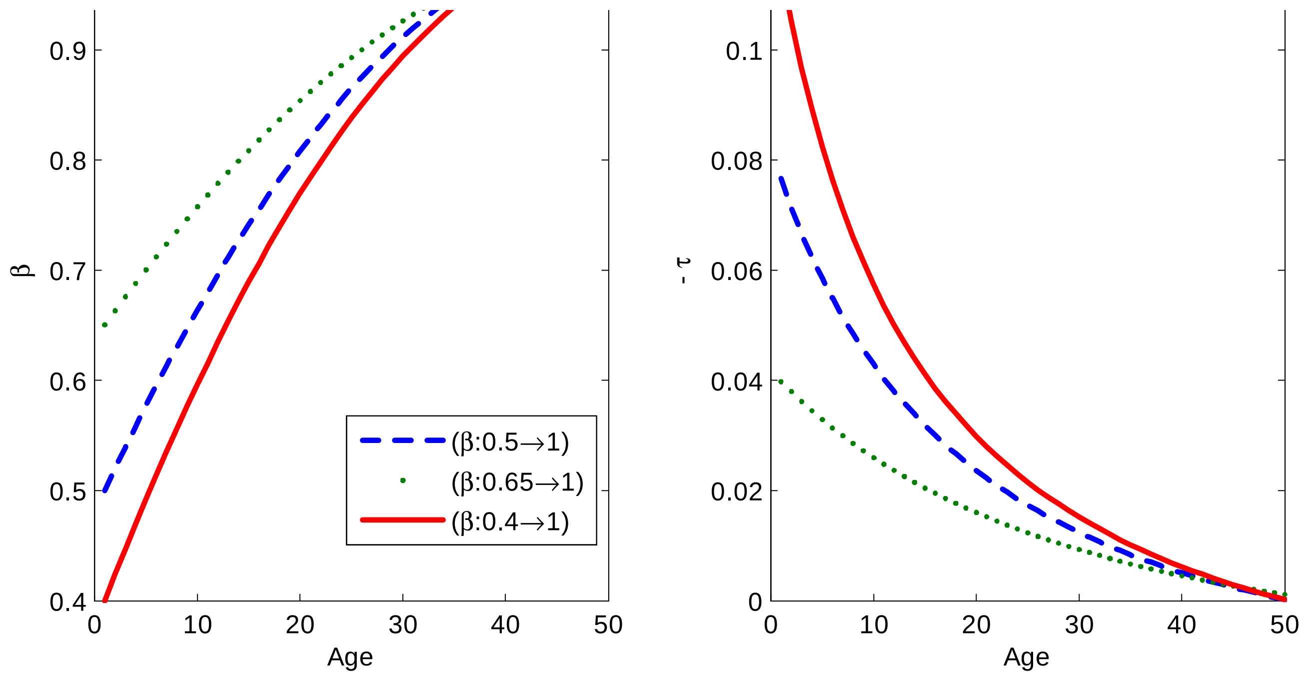

Figure 4: Sensitivity with respect to β_1

constant over the lifecycle. The blue solid curve in Figure 5 represents the benchmark case of full sophistication ($\pi = 1$) where optimal taxes do not depend on σ. We plot four other values of constant partial sophistication, $\pi = 0.3, 0.5, 0.7,$ and 0.9. It can be seen that as long as the level of partial sophistication is constant over the lifecycle, its effect on optimal taxes are negligible. Only at the very end, the tax curves for economies with partially sophisticated agents diverge from the benchmark and this is because from period 50 to 51 the level of sophistication jumps from its constant level to 1. To further scrutinize whether sudden jumps in partial sophistication from one period to another can cause significant divergence of the level of optimal capital subsidies from benchmark, we carry out experiments summarized in Figure 6 below. Here, the level of partial sophistication jumps twice over the lifecycle. First, it jumps temporarily from its constant level π to 1 in period 10 and then jumps back to its constant level. As we see this creates two jumps for optimal subsidies. The second occurs in period 25, but this is a permanent one: agent remains fully sophisticated from then on. Here, there is only one jump for optimal taxes since the level of partial sophsitication does not jump back to π. This, to us, is a hint that naivete matters only when there is a significant change in naivete in two consecutive periods. Finally, we analyze the effect of sophistication on optimal subsidies when the level of sophistication changes smoothly over the lifecycle. We assume π increases concavely. This experiment is summarized in Figure 7, where we see that when partial sophistication changes smoothly over the lifecycle, optimal capital subsidies are very similar to the benchmark case of full sophistication. We also do robustness checks for σ different from 2. As the figure 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

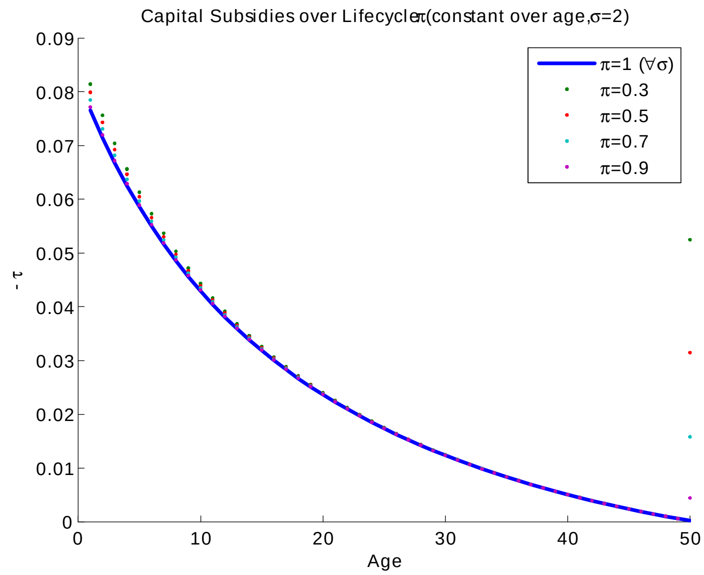

Figure 5: Sensitivity with respect to partial sophistication (constant π)

difference decreases to below 0.01% after the fourth period.

So, there are two major conclusions derived from the above set of experiments. First, as long as the level of naivete is not changing abruptly from one period to another, the level optimal capital subsidies over the lifecycle is robust to all scenarios about how sophistication changes with age. Second, as the last experiment shows, when we redo the experiments above with different σ in the range 0.5 to 5, we generalize the finding that optimal subsidies under partial sophistication are almost identical to their levels under full sophistication. Since under full sophistication σ is immaterial for optimal taxes, this means the level optimal capital subsidies over the lifecycle is not affected by our choice of the coefficient of constant relative risk aversion, σ, as long as the level of sophistication is not changing abruptly from one period to another.

We end this section by providing the following table which summarizes the effect of partial sophistication on bequest taxation. $\beta_1 = 0.5$ is assumed. First, we see that if $\pi_1 = 0.5$, then partial sophistication and σ matter for bequest taxes. This is in line with our previous numerical finding that partial sophistication only matters when there is a jump– here there is a jump in π from 1 to 0.5 (since parents are assumed to be fully sophisticated). When $\pi_1 = 0.95$, on the other hand, partial sophistication and/or σ do not matter much as the change in π is a much milder one.

σ	π_1	Bequest Tax
0.5	0.5	33%
0.5	0.95	45%
1	0.5	48%
1	0.95	48%
2	0.5	51%
2	0.95	48%

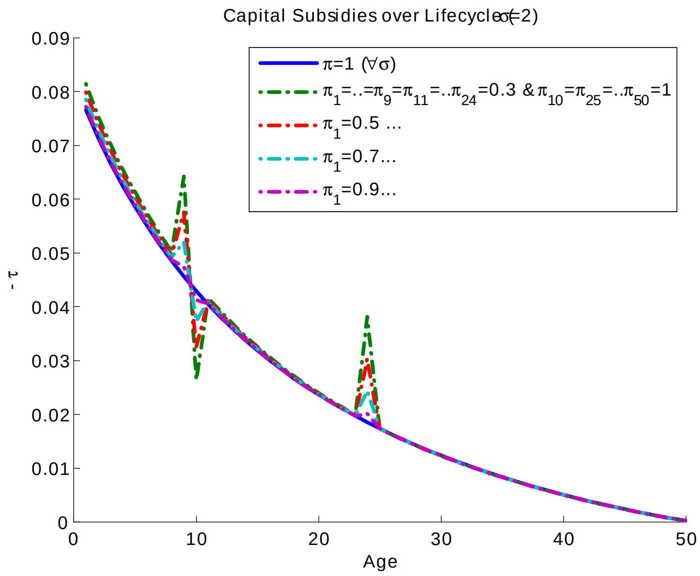

Figure 6: Sensitivity with respect to partial sophistication (jumps in π)

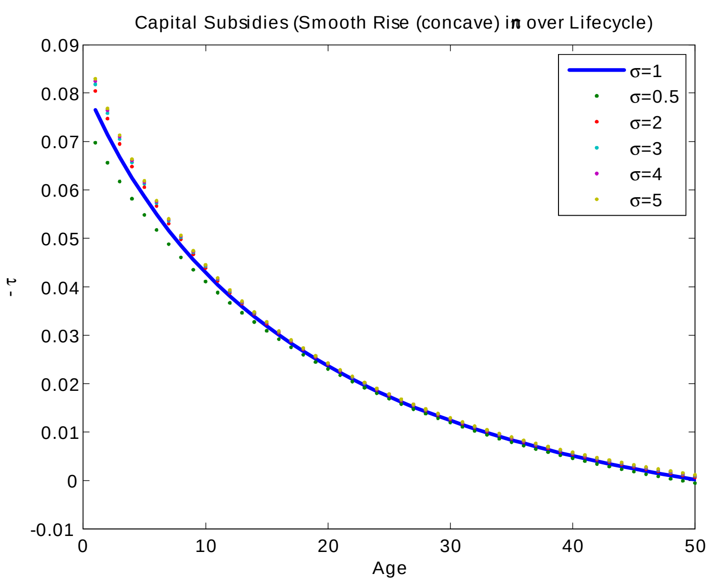

Figure 7: Sensitivity with respect to partial sophistication (smoothly rising π)

5.1 Intuition behind Correcting Naivete

In section 3, we saw that when utility is log the degree of sophistication of agents did not matter for the level of optimal taxes. Then, in section 4, we showed through numerical examples how optimal taxes depended on the level of agents' sophistication. In this subsection we want to understand the reason behind the numerical results. We will show analytically that if $\sigma > 1$ (< 1), then the optimal tax an agent pays increases (decreases) with her degree of sophistication. We will illustrate this through the simple three period example we have been using.

Self 1's problem:

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

Proposition 7 *If $\sigma > 1$ (< 1), then $\frac{\partial \tau_1^*}{\partial \pi_1} > 0 (< 0)$.*

Proof. *Relegated to Appendix A.* ■

The idea behind the proposition is as follows. Remember that the FOC wrt b_1 is:

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

When utility is log, $\left[1 + b_2'(b_1) \frac{\{-1+\frac{1}{\beta_2}\}}{R(1-\tau_1)} \right] = \frac{u'(c_2^N)}{u'(c_2)}$, and so self 1's tendency to oversave and hence taxes did not depend on the degree of sophistication. However, when $\sigma > 1$,

$$\left[1 + b_2'(b_1) \frac{\{-1 + \frac{1}{\beta_2}\}}{R(1-\tau_1)} \right] > \frac{u'(c_2^N)}{u'(c_2)},$$

which implies self 1's saving is actually higher when he is fully sophisticated than whenhe is fully naive. As a result, as the level of sophistication increases, his savings level increasesm which implies the tax he should pay should also increase. The story when $\sigma < 1$ is is identical.

6 Discussion of some assumptions

In this section, we want to relax two assumptions we made in the main body of the paper and show that our results do not hinge upon these assumptions.

6.1 Illiquid Assets

This subsection shows that introducing illiquid assets do not change the results on optimal taxes as long as there are no borrowing constraints. Specifically, through a 3 period example, we will show that the exact tax system that implements first-best in an environment without illiquid assets, also implements first-best in the environment with illiquid assets.

The illiquid asset d_1 does not pay in period 2 but pays in period 3 an after tax return $R^d(1-\tau^d)d_1$. Self2's problem then is:

$$
\begin{aligned}
c_2, c_3 &\in \arg\max_{c_2,c_3} u(c_2) + \hat{\beta}\delta u(c_3) \\
& \quad s.t. \\
c_2 + \frac{c_3}{R(1-\tau_2)} &\leq R(1-\tau_1)b_1 + T_1 + \frac{T_2}{R(1-\tau_2)} + \frac{R^d(1-\tau^d)d_1}{R(1-\tau_2)} = w_1(b_1, d_1)
\end{aligned}
$$

Let $c_2(w_1), c_3(w_1)$ be the solution to the above problem when $\hat{\beta} = \beta_2$ and $c_2^N(w_1), c_3^N(w_1)$ when $\hat{\beta} = 1$. Then, one can write the problem of self1 as:

$$
\begin{aligned}
&\max_{b_1,d_1} u(w - b_1 - d_1) + \pi_1 \delta \left[u(c_2(w_1)) + \delta u(c_3(w_1))\right] \\
&+(1-\pi_1)\delta \left[u(c_2^N(w_1)) + \delta u(c_3^N(w_1))\right]
\end{aligned}
$$

If government sets taxes such that

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

obviously, $d_1 = 0$. So, it is as if there are no illiquid assets; government prevents the usage of these assets through taxes. Then, simply by setting τ_1, τ_2 exactly according to first best taxes in the environment without illiquid asset, τ_1, τ_2, we get first best implemented in the market with illiquid asset. Let's compute these taxes for future use. Since

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

first-best requires

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

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

$$u'(c_1) = \delta\left(\pi_1\left[u'(c_2(w_1))c_2'(w_1)\frac{\partial w_1}{\partial b_1} + \delta u'(c_3(w_1))c_3'(w_1)\frac{\partial w_1}{\partial b_1}\right] + (1-\pi_1)\left[u'(c_2^N(w_1))c_2'^N(w_1)\frac{\partial w_1}{\partial b_1} + \delta u'(c_3^N(w_1))c_3'^N(w_1)\frac{\partial w_1}{\partial b_1}\right]\right)$$

where $\frac{\partial w_1}{\partial b_1} = R(1-\tau_1)$. So:

$$u'(c_1) = \delta R(1-\tau_1)\left(\pi_1\left[u'(c_2(w_1))c_2'(w_1) + \delta u'(c_3(w_1))c_3'(w_1)\right] + (1-\pi_1)\left[u'(c_2^N(w_1))c_2'^N(w_1) + \delta u'(c_3^N(w_1))c_3'^N(w_1)\right]\right)$$

which implies:

$$(1-\tau_1^*) = \frac{u'(c_1^*)}{\delta R\left(\pi_1\left[u'(c_2^*)c_2' + \delta u'(c_3^*)c_3'\right] + (1-\pi_1)\left[u'(c_2^{*N})c_2'^N + \delta u'(c_3^{*N})c_3'^N\right]\right)}.$$

Now suppose

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

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

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

still holds, first best still requires

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

To see period one taxes, consider the foc wrt d_1 :

$$u'(c_1) = \delta\left(\pi_1\left[u'(c_2(w_1))c_2'(w_1)\frac{\partial w_1}{\partial d_1} + \delta u'(c_3(w_1))c_3'(w_1)\frac{\partial w_1}{\partial d_1}\right] + (1-\pi_1)\left[u'(c_2^N(w_1))c_2^{\prime N}(w_1)\frac{\partial w_1}{\partial d_1} + \delta u'(c_3^N(w_1))c_3^{\prime N}(w_1)\frac{\partial w_1}{\partial d_1}\right]\right)$$

where $\frac{\partial w_1}{\partial b_1} = \frac{R^d(1-\tau^d)}{R(1-\tau_2)}$. So:

$$u'(c_1) = \delta\frac{R^d(1-\tau^d)}{R(1-\tau_2)}\left(\pi_1\left[u'(c_2(w_1))c_2'(w_1) + \delta u'(c_3(w_1))c_3'(w_1)\right] + (1-\pi_1)\left[u'(c_2^N(w_1))c_2^{\prime N}(w_1) + \delta u'(c_3^N(w_1))c_3^{\prime N}(w_1)\right]\right)$$

which implies:

$$\begin{aligned} R^d(1-\tau^d) &= R(1-\tau_2^*)\frac{u'(c_1^*)}{\delta R\left(\pi_1\left[u'(c_2^*)c_2' + \delta u'(c_3^*)c_3'\right] + (1-\pi_1)\left[u'(c_2^{*N})c_2^{\prime N} + \delta u'(c_3^{*N})c_3^{\prime N}\right]\right)} \\ &= R^2(1-\tau_2^*)R(1-\tau_1^*). \end{aligned}$$

Therefore, even if there is an illiquid asset, government can reach first best by setting taxes exactly equal to the first-best taxes in the environment without illiquid assets.

7 Conclusion

This paper studies optimal bequest and capital taxation in an economy where agents face self-control problems. In contrary to the previous literature and in line with evidence suggested by personality psychology and experimental studies we assume that the severity of the self-control problem changes over the life-cycle. We also allow for age-dependent partial sophistication. We restrict attention to CRRA utility functions and show that (1) if the utility function is further restricted to logarithmic function, then optimal taxes are independent of the pattern of partial sophistication; (2) if all the agents are assumed to be fully sophisticated and a steady-state condition holds, then optimal taxes are independent of CRRA coefficient. We derive explicit formulas which allow us to compute optimal taxes given the evolution of self-control problem over the life-cycle. In particular, 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. We then move away simultaneously from the assumptions of logarithmic utility and full sophistication to analyze the effects of naivete on optimal taxes. Using numerical simulations, we find that, as long as people's level of sophistication does not change abruptly from one period to another, optimal taxes we compute are robust to any pattern of partial sophistication over the life-span.

8 References

Ameriks, John, Andrew Caplin, John Leahy, and Tom Tyler. 2007. "Measuring Self-Control Problems." *American Economic Review,* 97(3): 966-72.

Ariely, Dan, and Klaus Wertenbroch. 2002. "Procrastination, Deadlines, and Performance: Self-Control by Precommitment." Psychological Science, 13(3): 219-24.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, Bas ter Weel. 2008. "The Economics and Psychology of Personality Traits," *Journal of Human Resources,* 43(4).

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature,* 47(2): 315-72.

Frederick, Shane, George Lowenstein, and Ted O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature,* 40(2): 351-401.

Green, Leonard, Astrid F. Fry, and Joel Myerson. 1994. "Discounting of Delayed Rewards: A Life-Span Comparison." Psychological Science, 5: 33-36.

Green, Leonard, Joel Myerson, Pawel Ostaszewski. 1999. "Discounting of Delayed Rewards across the Life-Span: Age Differences in Individual Discounting Functions." Behavioural Processes, 46: 89-96.

Gul, Faruk, Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica,* 69(6): 1403:35.

Helson, Ravenna, Constance Jones, and Virginia S. Y. Kwan. 2002. "Personality Change Over 40 Years of Adulthood: Hierarchical Linear mmodeling Analyses of Two Longitudinal Samples." *Journal of Personality and Social Psychology,* 83(3): 752-66.

Krusell, Per, Anthony Smith, and Burhanettin Kuruscu. 2002. "Equilibrium Welfare and Government Policy with Quasi-Geometric Discounting." *Journal of Economic Theory,* 105(1): 42-72.

Krusell, Per, Anthony Smith, and Burhanettin Kuruscu. 2009. "How Much Can Taxation Alleviate Temptation and Self-Control Problems?" Working Paper.

Krusell, Per, Anthony Smith, and Burhanettin Kuruscu. 2010. "Temptation and Taxation." *Econometrica,* 78(6): 2063-84.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics,* 112(2): 443-77.

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* 89(1): 103-24.

Phelps, E. S. and R. A. Pollak. 1968. "On Second-best National Saving and Game-Equilibrium Growth." *Review of Economics Studies,* (35): 185-199.

Read, Daniel, and N. L. Read. 2004. "Time Discounting over the Lifespan." *Organizational Behavior and Human Decision Processes,* 94(1): 22-32.

Roberts, Brent W., Richard W. Robins, Kali H. Trzesniewski, and Avshalom Caspi. 2003. "Personality Trait Development in Adulthood." *Handbook of Life Course,* ed. by Jeylan T. Mortimer and Michael J. Shanahan. New York.

Roberts, Brent W., Kate E. Walton, and Wolfgang 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.

Srivastava, Sanjay, Oliver P. John, Samuel D. Gosling, Jeff Potter. 2003. "Development of Personality in Early and Middle Adulthood: Set Like Plaster or Persistent Change?" *Journal of Personality and Social Psychology,* 84(5): 1041-53.

Strotz, Robert. 1955. "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies,* 23(3): 165-80.

Thaler, Richard H., and H. M. Shefrin. 1981. "An Economic Theory of Self-Control." *Journal of Political Economy,* 89(2): 392-406.

9 Appendix A - Proofs

9.1 Proof of Proposition 1.

Guess

$$V\left(b;\boldsymbol{\tau}\right)=D\log(\Gamma_I\left(b\right))+B$$

where D is a constant. Also,

$$W_i(b;\tau)=D_i\log(\Gamma_i\left(b\right))+B_i,$$

for $i=0,..,I$ where $D_0,D_1,..,D_I$ and $B_0,...,B_I$ are constants.

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

Now, from the FOC for the W_i problem, we have:

$$b_i\left(b\right)=\frac{R(1-\tau_i)b+T_i-\left[G_{i+1}+T_{i+1}\right]\left[\delta R(1-\tau_{i+1})D_{i+1}\right]^{-1}}{1+\left[\delta R(1-\tau_{i+1})D_{i+1}\right]^{-1}R(1-\tau_{i+1})}$$

Pluggin this in the value function, we get for $i=0,1,..,I$:

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

and

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

Thus,

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

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

Take $D_0,..D_I$ from above. Compute V',W_i' in terms of D,D_i. Plug these V',W_i' in the IC to get:

$$\begin{aligned}V'(b_I;\tau)&=D\Gamma_I'(b_I)(\Gamma_I(b_I))^{-1}=DR(1-\tau_I)(\Gamma_I(b)Q_I)^{-1}\\W_i'(b_i;\tau)&=D_i\Gamma_i'(b_i)(\Gamma_i(b_i))^{-1}=D_iR(1-\tau_i)(\Gamma_I(b)Q_i)^{-1}\end{aligned}$$

Plugging these and the fact that $\prod_{j=2}^{i}\left(R(1-\tau_j)(1-M_j)\right)=\frac{Q_i}{Q_1}$in the ICs, we get:

$$
\begin{aligned}
for\ all\ i &\in \{0,...,I-2\} \\
(M_{i+1}Q_i)^{-1} &= \delta\beta_{i+1}R(1-\tau_{i+1})\,(Q_{i+1})^{-1}\left[\begin{array}{c}\pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{j-(i+2)}+\delta^{I-(i+1)}D\right\} \\ +\,(1-\pi_{i+1})\,D_{i+1}\end{array}\right] && (11) \\
& and && (12) \\
(M_I Q_{I-1})^{-1} &= \delta\beta_I R(1-\tau_I)Q_I^{-1}\left[\pi_I D+(1-\pi_I)\,D_I\right] && \text{(self-3)}
\end{aligned}
$$

Parent's FOC:

$$
(M_0\Gamma_I(b))^{-1}=\delta R(1-\tau_0)\left[\sum_{i=0}^{I-1}\delta^i\left(M_{i+1}Q_i\Gamma_I(b)\right)^{-1}M_{i+1}\prod_{j=1}^{i}\left(R(1-\tau_j)(1-M_j)\right)+\delta^I V'(b_I;\boldsymbol{\tau})(1-M_I)\prod_{j=1}^{I-1}\left(R(1-\tau_j)(1-M_j)\right)\right]
$$

which after some simple algebra becomes:

$$
(M_0)^{-1}=\delta R(1-\tau_0)Q_0^{-1}\left[\sum_{i=0}^{I-1}\delta^i+\delta^I D\right].
$$

Now, using the last IC, 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 ICs to compute $M_{i+1}(D)$ for $i=0,..,I-2:$

$$
M_{i+1}(D)=\frac{1}{1+\beta_{i+1}\delta\left(\pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{j-(i+2)}+\delta^{I-(i+1)}D\right\}+(1-\pi_{i+1})\,D_{i+1}\right)}
$$

Now take these and plug them in the parent's problem and take first order condition w.r.t. b_0 to compute $M_0(D):$

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

Now verify the value function to get:

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

which implies

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

and hence

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

Now we turn to taxes that implement first-best level of bequest and savings.

The $I+1$ Euler Equations (of self1, self2, .., selfI and the parent) are restated below, these EEs are useful in computing FB taxes.

$$
\begin{aligned}
\left(M_1 Q_0 \Gamma_I(b)\right)^{-1} &= \delta\beta_1 R(1-\tau_1)\left(M_2 Q_1 \Gamma_I(b)\right)^{-1} \frac{\left[\begin{array}{c}\pi_1\left\{\sum_{i=2}^{I}\delta^{i-2}+\delta^{I-1}D\right\}\\ +\left(1-\pi_1\right)D_1\end{array}\right]}{M_2^{-1}} && (13)\\
&. \\
\left(M_{i+1} Q_i \Gamma_I(b)\right)^{-1} &= \delta\beta_{i+1} R(1-\tau_{i+1})\left(M_{i+2} Q_{i+1} \Gamma_I(b)\right)^{-1} \frac{\left[\begin{array}{c}\pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{j-(i+2)}+\delta^{I-(i+1)}D\right\}\\ +\left(1-\pi_{i+1}\right)D_{i+1}\end{array}\right]}{M_{i+2}^{-1}} && (14)\\
&. && (15)\\
\left(M_I Q_{I-1} \Gamma_I(b)\right)^{-1} &= \delta\beta_I R(1-\tau_I)\left(M_0 Q_I \Gamma_I(b)\right)^{-1} \frac{\left[\pi_I D+\left(1-\pi_I\right)D_I\right]}{M_0^{-1}} && (16)\\
\left(M_0 \Gamma_I(b)\right)^{-1} &= \delta R(1-\tau_0)\left(M_1 Q_0 \Gamma_I(b)\right)^{-1} \frac{\left[\sum_{i=0}^{I-1}\delta^{i}+\delta^{I}D\right]}{M_1^{-1}} && (17)
\end{aligned}
$$

Now comparison of above EEs with the corresponding FB EEs, we immediately see that:

$$
\begin{aligned}
\text{For all } i &\in \{0,..,I-2\},\\
\left(1-\tau_{i+1}^{*}\right) &= \frac{1}{\beta_{i+1}}\left(\frac{\left[\pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{j-(i+2)}+\delta^{I-(i+1)}D\right\}+\left(1-\pi_{i+1}\right)D_{i+1}\right]}{M_{i+2}^{-1}}\right)^{-1} = \frac{1}{\beta_{i+1}}\left(\frac{M_{i+2}^{-1}}{D}\right) = \frac{1}{\beta_{i+1}}\frac{1+\beta_{i+2}\delta D}{D} = \frac{1}{\beta_{i+1}}\left(1-\delta+\beta_{i}\right.\\
\left(1-\tau_I^{*}\right) &= \frac{1}{\beta_I}\frac{M_0^{-1}}{D} = \frac{1}{\beta_I}\frac{D}{D} = \frac{1}{\beta_I},\\
\left(1-\tau_0^{*}\right) &= \frac{M_1^{-1}}{D} = \frac{1+\beta_1\delta D}{D} = 1-\delta+\beta_1\delta.
\end{aligned}
$$

where we used the fact that plugging in $D_0 = .. = D_I = D = \frac{1}{1-\delta}$ gives:

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

Similarly, use other ICs to compute $M_i(D)$ for $i = 1,..,I-1$:

$$
M_{i+1}(D) = \frac{1}{1+\beta_{i+1}\delta D}
$$

Now take these and plug them in the parent's problem and take first order condition w.r.t. b_0 to compute $M_0(D) = \frac{1}{D}$.

9.2 Proof of Proposition 2.

Take the incentive constraint for agent i. It is straightforward but messy to show that if we plug in the incentive constraints for $j > i$ in $i's$ incentive constraint, we get:

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

which renders first-best taxes as:

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

Now suppose utility is CRRA and policies are linear in current wealth. This implies:

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

Now plug this in the general formula for tax to get the CRRA specific tax formula:

$$(1-\tau_i^*) = \frac{1}{\beta_i}\frac{1}{1+(1-M_{i+1}^*)\left(\frac{1}{\beta_{i+1}}-1\right)}. \tag{20}$$

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

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

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

$$M_i^* = \frac{M_{i+1}^*R(1-\tau_i^*)}{1+M_{i+1}^*R(1-\tau_i^*)}. \tag{21}$$

Plugging (20) in (21), we get a system of $(I+1)$ equations in $(I+1)$ unknows $(M_0^*, ..., M_I^*)$ that fully pin down agents policies when they face first-best taxes, for the CRRA 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 log solution,

$$M_i = \frac{1-\delta}{1-\delta+\beta_i\delta} \tag{22}$$

for all i, satisfies the above system of equations. So, when $R\delta = 1$, policies are given by (22). Plugging (22) in (20), the formula for taxes, we get the first-best taxes

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

which is exactly the same as (**??**) once we introduce $\beta_0 = 1$.

9.3 Proof of Proposition 7.

The FOC of self1 reads:

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

When utility is CRRA, a self2 that has self control problem level $\hat{\beta}_2$ consumes according to:

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

where

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

With probability π_1 self1 correctly believes that $\hat{\beta}_2 = \beta_2$ and with probability $1-\pi_1$, he believes that self2 has no self control problems and hence $\hat{\beta}_2 = 1$.

Now it follows from the budget that:

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

which implies

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

Plugging this in the FOC, we get:

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

where with some abuse of notation $\hat{c}_2, \hat{c}_3$ denote consumption levels when $\hat{\beta}_2 = 1$.

So, the first best tax is:

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

where

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

Remember we want to compute:

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

since first-best allocation is independent of π_1.

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

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

Then,

$$\begin{aligned} \frac{\partial \frac{\kappa(\hat{\beta}_2)}{\kappa(\beta_2)}}{\partial \hat{\beta}_2} &= \kappa(\beta_2)^{-1} \left[\begin{array}{c} (1-\sigma) \left(\frac{R(1-\tau_2)}{R(1-\tau_2)+\left(R(1-\tau_2)\hat{\beta}_2\delta\right)^{\frac{1}{\sigma}}} \right)^{-\sigma} R(1-\tau_2)(-1) \left(\frac{1}{R(1-\tau_2)+\left(R(1-\tau_2)\hat{\beta}_2\delta\right)^{\frac{1}{\sigma}}} \right)^2 \\ \frac{1}{\sigma} \left(R(1-\tau_2)\hat{\beta}_2\delta \right)^{\frac{1}{\sigma}-1} R(1-\tau_2)\delta \left\{ 1 + \delta \left(R(1-\tau_2)\hat{\beta}_2\delta \right)^{\frac{1}{\sigma}-1} \right\} \\ + \left(\frac{R(1-\tau_2)}{R(1-\tau_2)+\left(R(1-\tau_2)\hat{\beta}\delta\right)^{\frac{1}{\sigma}}} \right)^{1-\sigma} \delta \left(\frac{1}{\sigma} - 1 \right) \left(R(1-\tau_2)\hat{\beta}\delta \right)^{\frac{1}{\sigma}-2} R(1-\tau_2)\delta \end{array} \right] \\ &= \kappa(\beta_2)^{-1} \left[\begin{array}{c} \frac{(1-\sigma)}{\sigma} \left(\frac{R(1-\tau_2)}{R(1-\tau_2)+\left(R(1-\tau_2)\hat{\beta}_2\delta\right)^{\frac{1}{\sigma}}} \right)^{1-\sigma} R(1-\tau_2)\delta \left(R(1-\tau_2)\hat{\beta}_2\delta \right)^{\frac{1}{\sigma}-1} \\ \left(\frac{1}{\hat{\beta}_2} - 1 \right) \left\{ \frac{1}{R(1-\tau_2)\left(R(1-\tau_2)+\left(R(1-\tau_2)\hat{\beta}_2\delta\right)^{\frac{1}{\sigma}} \right)} \right\} \end{array} \right] \end{aligned}$$

which implies

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

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

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

10 Appendix B - Computational Procedure

10.1 Guess:

Guess

$$V(b; \boldsymbol{\tau}) = D(\tau) \frac{(\Gamma_3(b))^{1-\sigma}}{1-\sigma}$$

where D is the constant of the parent's value function which depends on the tax system τ. Also,

$$W_i(b; \tau) = D_i(\tau) \frac{(\Gamma_i(b))^{1-\sigma}}{1-\sigma},$$

where D_i for $i = 0, 1, .., I$ is the constant of fully naive self i's value function.

10.2 Computing Equilibrium for a given tax system τ (not necessarily optimal):

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

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

$$D_i = \left[\frac{[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1}{\sigma}} R(1-\tau_{i+1})}{1+[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1}{\sigma}} R(1-\tau_{i+1})}\right]^{1-\sigma} \left(1+\delta\frac{D_{i+1}}{[\delta R(1-\tau_{i+1})D_{i+1}]^{-\frac{1-\sigma}{\sigma}}}\right)$$

and

$$D_I = \left[\frac{[\delta R(1-\tau_0)D_0]^{-\frac{1}{\sigma}} R(1-\tau_0)}{1+[\delta R(1-\tau_0)D_0]^{-\frac{1}{\sigma}} R(1-\tau_0)}\right]^{1-\sigma} \left(1+\delta\frac{D_0}{[\delta R(1-\tau_0)D_0]^{-\frac{1-\sigma}{\sigma}}}\right).$$

Then, we just plug in the given tax values from τ and then solve a $I+1$ equations and $I+1$ unknowns.

STEP 2: Computing equilibrium D.

Take $D_0,..,D_I$ from STEP 1 above. Compute V', W_i' for all i in terms of D, D_i :

$$\begin{aligned} V'(b_I;\tau) &= D\Gamma_I'(b_I)(\Gamma_I(b_I))^{-\sigma} = DR(1-\tau_I)(\Gamma_I(b)Q_I)^{-\sigma} \\ W_i'(b_i;\tau) &= D_i\Gamma_i'(b_i)(\Gamma_i(b_i))^{-\sigma} = D_iR(1-\tau_i)(\Gamma_I(b)Q_i)^{-\sigma} \end{aligned}$$

Plug these V', W_i' in the ICs to transform them to:

$$\text{For all } i \in \{0,..,I-2\},\ (M_{i+1}Q_i)^{-\sigma} = \delta\beta_{i+1}R(1-\tau_{i+1})(Q_{i+1})^{-\sigma} \left[\begin{array}{c} \pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{j-(i+2)}\left(M_j\frac{Q_{j-1}}{Q_{i+1}}\right)^{1-\sigma} + \delta^{I-(i+1)}D\left(\frac{Q_I}{Q_{i+1}}\right)^{1-\sigma}\right\} \\ +(1-\pi_{i+1})D_{i+1} \end{array}\right]$$

$$(M_IQ_{I-1})^{-\sigma} = \delta\beta_IR(1-\tau_I)Q_I^{-\sigma}\left[\pi_ID+(1-\pi_I)D_I\right]$$

Using the last IC, it is easy to show that

$$M_I(D) = \frac{1}{1+\left(\beta_I\delta\left[R(1-\tau_I)\right]^{1-\sigma}(\pi_ID+(1-\pi_I)D_I)\right)^{\frac{1}{\sigma}}}.$$

Then, use other ICs to compute $M_1(D), ..M_{I-1}(D)$ recursively as follows:

$$\text{For all } i \in \{0,..,I-2\},\ M_{i+1} = \frac{1}{1+\left\{\delta\beta_{i+1}\left[R(1-\tau_{i+1})\right]^{1-\sigma}\left[\begin{array}{c} \pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{j-(i+2)}\left(M_j\frac{Q_{j-1}}{Q_{i+1}}\right)^{1-\sigma} + \delta^{I-(i+1)}D\left(\frac{Q_I}{Q_{i+1}}\right)^{1-\sigma}\right\} \\ +(1-\pi_{i+1})D_{i+1} \end{array}\right]\right\}^{\frac{1}{\sigma}}}$$

Now plug these in the parent's problem and take first order condition w.r.t. M_0:

$$\begin{aligned} &\max_{M_0,..,M_I} \frac{(M_0\Gamma_I(b))^{1-\sigma}}{1-\sigma} + \delta\left[\sum_{i=1}^{I-1}\delta^i\frac{(Q_{i-1}M_i\Gamma_I(b))^{1-\sigma}}{1-\sigma} + \delta^ID\frac{(\Gamma_I(b)Q_I)^{1-\sigma}}{1-\sigma}\right] \\ \equiv &\max_{M_0} \frac{(M_0\Gamma_I(b))^{1-\sigma}}{1-\sigma} + \delta\left[\sum_{i=1}^{I-1}\delta^i\frac{(Q_{i-1}M_i(D)\Gamma_I(b))^{1-\sigma}}{1-\sigma} + \delta^ID\frac{(\Gamma_I(b)Q_I)^{1-\sigma}}{1-\sigma}\right] \\ \equiv &\max_{M_0} \frac{(M_0)^{1-\sigma}}{1-\sigma} + \delta\frac{Q_0^{1-\sigma}}{1-\sigma}\left[\sum_{i=1}^{I-1}\delta^i\left(\frac{Q_{i-1}}{Q_0}M_i(D)\right)^{1-\sigma} + \delta^ID\left(\frac{Q_I}{Q_0}\right)^{1-\sigma}\right] \end{aligned}$$

Call $A(D) \equiv \left[\sum_{i=1}^{I-1}\delta^i\left(\frac{Q_{i-1}}{Q_0}M_i(D)\right)^{1-\sigma} + \delta^ID\left(\frac{Q_I}{Q_0}\right)^{1-\sigma}\right]$.

Taking the FOC:

$$M_0^{-\sigma} + \delta Q_0^{-\sigma} \frac{\partial Q_0}{\partial M_0} A(D) = 0$$

which gives M_0 as a function of D :

$$M_0^{-\sigma} = \delta \left[R(1-\tau_0)\right]^{1-\sigma} (1-M_0)^{-\sigma} A(D),$$

which gives:

$$M_0(D) = \frac{1}{\left(\delta \left[R(1-\tau_0)\right]^{1-\sigma} A(D)\right)^{\frac{1}{\sigma}} + 1}.$$

Plug this in the parent's problem to get D' as a function of D :

$$D'(D) = M_0(D)^{1-\sigma} + \delta \left[R(1-\tau_0)(1-M_0(D))\right]^{1-\sigma} A(D).$$

Then, we look for the fixed points of D.

Once we compute D, we can first compute $M_I(D)$, then we compute $M_{I-1}(D)$, and so on until $M_0(D)$.

10.3 Computing Optimal Tax System $\boldsymbol{\tau}^*$:

The $I+1$ Euler Equations of self 1 through self I and the parent are restated below.

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

Note that envelope suggests that $D = M_0^{-\sigma}$ as from the closed form $V'(b) = D\left[R(1-\tau_3)\right]^{1-\sigma}$ and by envelope $V'(b) = R(1-\tau_3) u'(c) = R(1-\tau_3)\left(M_0 R(1-\tau_3)\right)^{-\sigma}$.

Now comparison of above EEs with the corresponding FB EEs, we immediately see that:

$$
\begin{aligned}
\text{For all } i &\in \{0,..,I-2\}, \qquad (24)\\
(1-\tau^*_{i+1}) &= \frac{1}{\beta_{i+1}}\left(\frac{\left[\pi_{i+1}\left\{\sum_{j=i+2}^{I}\delta^{i-(i+2)}\left(M^*_j\frac{Q^*_{j-1}}{Q^*_{i+1}}\right)^{1-\sigma}+\delta^{I-(i+1)}D^*\left(\frac{Q^*_I}{Q^*_{i+1}}\right)^{1-\sigma}+(1-\pi_{i+1})\,D^*_{i+1}\right\}\right]}{M^{*-\sigma}_{i+2}}\right)^{-1}\\
(1-\tau^*_I) &= \frac{1}{\beta_I}\left(\frac{[\pi_I D^*+(1-\pi_I)\,D^*_I]}{M^{*-\sigma}_0}\right)^{-1}\\
(1-\tau^*_0) &= \left(\frac{\left[\sum_{i=1}^{I}\delta^{i-1}\left(M_i\frac{Q^*_{i-1}}{Q^*_0}\right)^{1-\sigma}+\delta^I D^*\left(\frac{Q^*_I}{Q^*_0}\right)^{1-\sigma}\right]}{M^{*-\sigma}_1}\right)^{-1},
\end{aligned}
$$

where D^*_i are FB values computed according to: for all $i \in \{0,1,..,I-1\}$

$$D^*_i = \left[\frac{\left[\delta R(1-\tau^*_{i+1})D^*_{i+1}\right]^{-\frac{1}{\sigma}}R(1-\tau^*_{i+1})}{1+\left[\delta R(1-\tau^*_{i+1})D^*_{i+1}\right]^{-\frac{1}{\sigma}}R(1-\tau^*_{i+1})}\right]^{1-\sigma}\left(1+\delta\frac{D^*_{i+1}}{\left[\delta R(1-\tau^*_{i+1})D^*_{i+1}\right]^{-\frac{1-\sigma}{\sigma}}}\right)$$

and

$$D^*_I = \left[\frac{[\delta R(1-\tau^*_0)D^*_0]^{-\frac{1}{\sigma}}R(1-\tau^*_0)}{1+[\delta R(1-\tau^*_0)D^*_0]^{-\frac{1}{\sigma}}R(1-\tau^*_0)}\right]^{1-\sigma}\left(1+\delta\frac{D^*_0}{[\delta R(1-\tau^*_0)D^*_0]^{-\frac{1-\sigma}{\sigma}}}\right).$$

This is a system of $2*(I+1)$ equations and $2*(I+1)$ unknowns (τ_i and D_i for $i=0,1,..,I$).

The only thing left before we go and solve these equations is to plug in the values of M^*_i in the first set of I+1 equations (23) using the formulas below:

$$
\begin{aligned}
M^*_0 &= \frac{c^*_0}{Rb(1-\tau^*_I)+T^*_I+G^*_I}=\frac{c^*_0}{Rb+G^*_I} \qquad (25)\\
\text{For all } i &\in \{0,1,..,I-1\},\ M^*_{i+1}=\frac{c^*_{i+1}}{Rb^*_i(1-\tau^*_i)+T^*_i+G^*_i}=\frac{c^*_{i+1}}{Rb^*_i+G^*_i}
\end{aligned}
$$

where

$$G^*_I = \frac{1}{1-\left[R^{I+1}\prod_{j=0}^{I}(1-\tau^*_j)\right]^{-1}}\sum_{i=0}^{I}\frac{T^*_i}{R^{i+1}\prod_{j=0}^{i}(1-\tau^*_j)}$$

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

$$G^*_i = \frac{G^*_{i+1}+Rb^*_{i+1}\tau^*_{i+1}}{R(1-\tau^*_{i+1})},$$

and $(c^*_i, b^*_i)^I_{i=0}$ are FB consumption and savings levels which can be computed relatively easily. The formula for these are:

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