

Optimal Life-cycle Capital Taxation under Self-Control Problems

ONLINE APPENDIX

Nicola Pavoni* and Hakki Yazici[†]

March 2015

A Proofs

A.1 Proof of Proposition 1.

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

Preparation to the proof.

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

*Bocconi University, IGIER, IFS, and CEPR.

[†]Sabanci University.

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

$$\begin{aligned}
V(b, k; \tau) &= \max_{b_0} u(R(k)(1 - \tau_1)b + w(k) + T_1 - b_0) \\
&\quad + \delta \left[\sum_{i=0}^{I-1} \delta^i u \left(R(k^{i+1})(1 - \tau_i) b_i(\dots(b_0) \dots) + w(k^{i+1}) + T_i - b_{i+1} \right) + \delta^I V \left(b_I(\dots(b_0) \dots), k^{I+1}; \tau \right) \right] \\
&\quad \text{s.t. for all } b_0 \\
b_1(b_0, k^1; \tau) &= \arg \max_{\hat{b}_1} u \left(R(k^1)(1 - \tau_0) b_0 + w(k^1) + T_0 - \hat{b}_1 \right) \\
+ \delta \beta_1 \left[\sum_{i=1}^{I-1} \delta^{i-1} u \left(R(k^{i+1})(1 - \tau_i) b_i(\dots(\hat{b}_1) \dots) + w(k^{i+1}) + T_i - b_{i+1}(\dots(\hat{b}_1) \dots) \right) + \delta^{I-1} V \left(b_I(\dots(\hat{b}_1) \dots), k^{I+1}; \tau \right) \right] \\
&\quad \text{s.t. for all } b_1 \\
b_2(b_1, k^2; \tau) &= \arg \max_{\hat{b}_2} u \left(R(k^2)(1 - \tau_1) b_1 + w(k^2) + T_1 - \hat{b}_2 \right) \\
+ \delta \beta_2 \left[\sum_{i=2}^{I-1} \delta^{i-2} u \left(R(k^{i+1})(1 - \tau_i) b_i(\dots(\hat{b}_2) \dots) + w(k^{i+1}) + T_i - b_{i+1}(\dots(\hat{b}_2) \dots) \right) + \delta^{I-2} V \left(b_I(\dots(\hat{b}_2) \dots), k^{I+1}; \tau \right) \right] \\
&\quad \text{s.t. for all } b_2 \\
&\quad \dots \\
&\quad \text{s.t. for all } b_{I-2} \\
b_{I-1}(b_{I-2}, k^{I-1}; \tau) &\in \arg \max_{\hat{b}_{I-1}} u \left(R(k^{I-1})(1 - \tau_{I-2}) b_{I-2} + w(k^{I-1}) + T_{I-2} - \hat{b}_{I-1} \right) \\
+ \delta \beta_{I-1} \left[u \left(R(k^I)(1 - \tau_{I-1}) \hat{b}_{I-1} + w(k^I) + T_{I-1} - b_I(\dots(\hat{b}_{I-1}) \dots) \right) + \delta V \left(b_I(\dots(\hat{b}_{I-1}) \dots), k^{I+1}; \tau \right) \right] \\
&\quad \text{s.t. for all } b_{I-1} \\
b_I(b_{I-1}, k^I; \tau) &= \arg \max_{\hat{b}_I} u \left(R(1 - \tau_{I-1}) b_{I-1} + w_{I-1} + T_{I-1} - \hat{b}_I \right) + \delta \beta_I V \left(\hat{b}_I, k^{I+1}; \tau \right),
\end{aligned}$$

where $b_i(\dots(b_i)\dots) = b_i$.

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

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

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 $i = 1, \dots, I$

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

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:

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

where $b_i(\dots(b)\dots)$ 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 $k^{s+1} = K(\dots(k)\dots)$, 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:

$$\begin{aligned}
V(b, k; \tau) &= \max_{M_0} u(M_0 \Gamma_I(b)) + \delta \left[\sum_{i=0}^{I-1} \delta^i u(Q_i(k) M_{i+1} \Gamma_I(b)) + \delta^I V\left((1 - M_I) Q_{I-1}(k) \Gamma_I(b), k^{I+1}; \tau\right) \right] \quad (2) \\
&\text{s.t. for all } i \in \{1, \dots, I-1\} \\
(M_i Q_{i-1}(k) \Gamma_I(b, k))^{-\sigma} &= \delta \beta_i \left[R(k^{i+1}) (1 - \tau_i(k^{i+1})) \left\{ \sum_{j=i+1}^I \delta^{j-(i+1)} (M_j Q_{j-1}(k) \Gamma_I(b, k))^{-\sigma} M_j \frac{Q_{j-1}(k)}{Q_i(k)} \right. \right. \\
&\quad \left. \left. + \delta^{I-i} V'(b_I(..(b)..), k^{I+1}; \tau) (1 - M_I) \frac{Q_{I-1}(k)}{Q_i(k)} \right\} \right] \\
(M_I Q_{I-1}(k) \Gamma_I(b, k))^{-\sigma} &= \delta \beta_I V'(b_I(..(b)..), k^{I+1}; \tau).
\end{aligned}$$

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

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

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

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

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

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

where we used the recursion (1).

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

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

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

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

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

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

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

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

Now, we verify the value function to compute D :

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

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

$$M_i = \frac{1 - \delta}{1 - \delta + \beta_i \delta}, \text{ for all } i \in \{1, \dots, I\}, \quad (4)$$

$$M_0 = 1 - \delta.$$

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 derivative of the value function from (3) :

$$(M_i Q_{i-1}(k) \Gamma_I(b, k))^{-1} = \delta \beta_i R(k^{i+1}) (1 - \tau_i(k^{i+1})) (M_{i+1} Q_i(k) \Gamma_I(b, k))^{-1} \left[\sum_{j=i+1}^I \delta^{j-(i+1)} + \delta^{I-i} D \right] M_{i+1}. \quad (5)$$

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

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

For self- I , the constraint describing his behavior in problem (2) 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

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

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:

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

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

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

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

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

Plugging (6) in (7), 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 (4) satisfies the above system of equations, meaning it is an equilibrium. Plugging (4) in the formula for taxes, (6), we get that optimal taxes are the same as the logarithmic utility case.

A.3 Proof of Proposition 6.

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

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

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

Guess

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

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 (1).

Plugging these in the constraints described in problem (2), 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 $i = 1, \dots, I-1$:

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

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

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

Now verify the value function to compute D :

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

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 (3) :

$$(M_i Q_{i-1}(k) \Gamma_I(b, k))^{-1} = \delta \beta_i R(k^{i+1}) (1 - \tau_i(k^{i+1})) (M_{i+1} Q_i(k) \Gamma_I(b, k))^{-1} \left[\sum_{j=i+1}^I \delta^{j-(i+1)} + \delta^{I-i} D \right] M_{i+1}. \quad (10)$$

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

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

which, using (9), implies

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

For self- I , the constraint describing his behavior in problem (2) 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

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

with the corresponding optimality condition gives

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

B Approximating Hyperbolic Discount Functions with Quasi-hyperbolic Discount Functions

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

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

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

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

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

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

Table 1: Approximating β from Hyperbolic Discount Functions

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

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

C Partial Sophistication

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

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

¹We 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)).

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

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

s.t. for all b_0

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

s.t. for all b_1

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

s.t. for all b_2

...

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

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

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:

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

with

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

To understand the nested nature of policies and the way we model partial sophistication better, let us analyze the definition of policies in (12) and (13). First, constraint (13) describes how self I chooses b_I . The number $\pi_I \in [0, 1]$ represents the belief of self I about the presence of self-control problems. More precisely, this is the belief of self I about the probability that next period when he becomes a parent he will face an offspring with self-control problems, i.e. $(\beta_1, \dots, \beta_I) \neq (1, \dots, 1)$, 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.² Second, consider constraint (12) which defines how self $I - 1$ chooses b_{I-1} . The number $\pi_{I-1} \in [0, 1]$ represents the degree of sophistication of self $I - 1$, meaning self $I - 1$ knows the truth that his followers will have self-control problems with probability π_{I-1} . In particular, with π_{I-1} probability self $I - 1$ thinks self I chooses b_I according to (13), and with the remaining probability he thinks self I chooses b_I without facing any self-control problems. We have just seen that the last constraint, (13), enters the parent's problem in at least two ways: first, in the definition of self I 's policy function and then as a constraint in the definition of self $I - 1$'s policy function. These two different constraints are represented by a single constraint, (13), because the parent and self $I - 1$'s sophisticated belief agree about how self I will behave.³ Similarly, the constraint describing self $I - 1$'s policy is also a constraint in the constraint that describes self $I - 2$'s policy, and self $I - 2$'s policy is also a constraint of self $I - 3$'s, and so on. Thus, actually the constraint that describes the policy of self i enters parent's problem in i different places but since these are all identical constraints, we represent them with just one constraint that describes self i 's policy.

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

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

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

Proof. Relegated to Appendix C.1. ■

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

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

²Of 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.

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

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

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

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

We conclude that, as long as the level of 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. Moreover, when the level of partial sophistication is changing smoothly (or constant), the level optimal capital subsidies over the life cycle is not significantly affected by our choice of the coefficient of constant relative risk aversion.

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

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

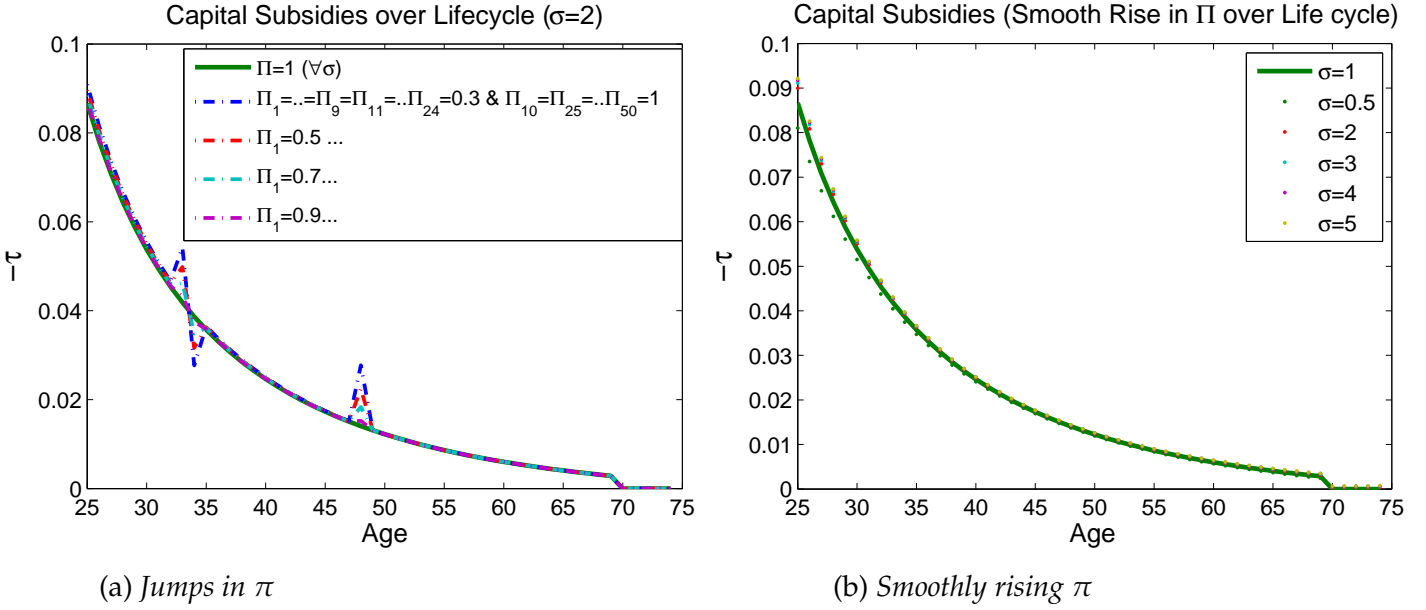


Figure 1: *Partial Sophistication*

C.1 Proof of Proposition C.1.

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

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

$$\begin{aligned}
 V(b, k; \tau) &= \max_{M_0} u(M_0\Gamma_I(b)) + \delta \left[\sum_{i=0}^{I-1} \delta^i u(Q_i(k)M_{i+1}\Gamma_I(b)) + \delta^I V\left((1 - M_I)Q_{I-1}(k)\Gamma_I(b), k^{I+1}; \tau\right) \right] \quad (14) \\
 &\text{s.t. for all } i \in \{1, \dots, I - 1\} \\
 (M_i Q_{i-1}(k)\Gamma_I(b, k))^{-\sigma} &= \delta \beta_i \left[\pi_i R(k^{i+1})(1 - \tau_i(k^{i+1})) \left\{ \sum_{j=i+1}^I \delta^{j-(i+1)} (M_j Q_{j-1}(k)\Gamma_I(b, k))^{-\sigma} M_j \frac{Q_{j-1}(k)}{Q_i(k)} \right. \right. \\
 &\quad \left. \left. + \delta^{I-i} V'(b_I(\dots(b)\dots), k^{I+1}; \tau)(1 - M_I) \frac{Q_{I-1}(k)}{Q_i(k)} \right\} \right. \\
 &\quad \left. + (1 - \pi_i) W'_i(b_i(\dots(b)\dots), k^{i+1}; \tau) \right] \\
 (M_I Q_{I-1}(k)\Gamma_I(b, k))^{-\sigma} &= \delta \beta_I \left[\pi_I V'(b_I(\dots(b)\dots), k^{I+1}; \tau) + (1 - \pi_I) W'_I(b_I(\dots(b)\dots), k^{I+1}; \tau) \right].
 \end{aligned}$$

Core proof of Proposition C.1.

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

the taxes specified by Proposition 1, people choose the efficient allocation.

Guess

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

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:

$$\begin{aligned} V'(b_I(\dots(b)\dots), k^{I+1}; \tau) &= DR(k^{I+1})(1 - \tau_I(k^{I+1}))(\Gamma_I(b, k)Q_I(k))^{-1}, \\ W'_i(b_i(\dots(b)\dots), k^{i+1}; \tau) &= D_i R(k^{i+1})(1 - \tau_i(k^{i+1}))(\Gamma_i(b, k)Q_i(k))^{-1}, \end{aligned} \quad (15)$$

where we used the recursion (1).

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

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

and

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

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 (14) and plugging in the $M_i(D)$ from above, we get:

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

Now verify the value function to compute D :

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

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

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

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

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

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

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

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

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

and the comparison of this with the efficiency condition gives

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

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

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

with the corresponding optimality condition gives

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

C.2 Computational Procedure

C.2.1 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.

C.2.2 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

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

C.2.3 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:

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

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

C.2.4 Iteration

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

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

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:

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

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

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.

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:

$$c_2, c_3 \in \arg \max_{c_2, c_3} u(c_2) + \bar{\beta}_2 \delta u(c_3)$$

$$\text{s.t.}$$

$$c_2 + \frac{c_3}{R(1 - \tau_2)} \leq R(1 - \tau_1)b_1 + T_1 + \frac{T_2}{R(1 - \tau_2)} + \frac{R^d(1 - \tau^d)d_1}{R(1 - \tau_2)} \equiv y_1(b_1, d_1)$$

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:

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

Case 1. Government sets taxes such that

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

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

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

efficiency requires

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

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

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

where $\frac{\partial y_1(b_1, d_1)}{\partial b_1} = R(1 - \tau_1)$ (For ease of exposition, assume the policies are differentiable).⁶ Therefore,

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

which implies:

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

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

Case 2. Government sets taxes such that

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

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

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

⁶It is well-known that in general we cannot guarantee even the continuity of the policy functions (e.g., see Krusell and Smith (2003), and Harris and Laibson (2001)).

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

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

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.

References

- ALI, S. N. (2011): "Learning Self-Control," *The Quarterly Journal of Economics*, 126(2), 857–893.
- ASHEIM, G. B. (2007): "Procrastination, Partial Naivete, and Behavioral Welfare Analysis," Memorandum 02/2007, Oslo University, Department of Economics.
- ELIAZ, K., AND R. SPIEGLER (2006): "Contracting with Diversely Naive Agents," *Review of Economic Studies*, 73(3), 689–714.
- 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.
- HARRIS, C. J., AND D. LAIBSON (2001): "Dynamic Choices of Hyperbolic Consumers," *Econometrica*, 69(4), 935–957.

- KRUSELL, P., AND A. A. SMITH (2003): "Consumption–Savings Decisions with Quasi-Geometric Discounting," *Econometrica*, 71(1), 365–375.
- 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.
- (2003): "Self Awareness and Self Control," in *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, pp. 217–243. Russell Sage Foundation Press.
- POLLAK, R. A. (1968): "Consistent Planning," *The Review of Economic Studies*, 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.