Economics 2450A: Public Economics
Section 9: Linear Capital Taxation

Matteo Paradisi

November 7, 2016

In this section we introduce a framework to study optimal linear capital taxation. We first focus on a two-period model, define the concept of intertemporal wedge and derive optimal capital taxes using the Atkinson Stiglitz result. We then move to an infinite horizon model with aggregate uncertainty and derive optimal taxes. Finally, we study a model with capitalists and workers and show that only under some assumption about preferences we can recover a zero capital tax in steady state.[1]

1 A Two-Period Model

We introduce a two-period model with capital accumulation that will be useful to study the problem of capital taxation. The two time periods are $t = 0$ and $t = 1$. The preferences for individual i are $U^i(c_0, c_1, y_0)$. The individual can save period 0 income and earn gross interest rate R on savings. We start by constraining the government's instruments and focusing on a linear consumption tax, while keeping a non-linear tax on income. The budget constraints for the two periods read:

$$(1 + \tau_0) c_0 \leq y_0 - T(y_0) - k_1$$

$$(1 + \tau_1) c_1 \leq Rk_1$$

Combining the two we get:

$$(1 + \tau_1) c_1 \leq R(y_0 - T(y_0) - (1 + \tau_0) c_0) \tag{1}$$

Rearranging we have:

$$c_0 + \frac{1}{R}\left[\frac{1 + \tau_1}{1 + \tau_0}\right] c_1 \leq \frac{y_0 - T(y_0)}{1 + \tau_0}$$

Let us denote $1 + \tilde{\tau}_1 = (1 + \tau_1) / (1 + \tau_0)$ and $\tilde{T}(y) = (\tau_0 y + T(y)) / (1 + \tau_0)$. The budget constraint can be written:

$$c_0 + \frac{1}{R}[1 + \tilde{\tau}_1] c_1 \leq y_0 - \tilde{T}(y_0)$$

When the budget constraint holds with equality we can write consumption at time 1 as:

$$c_1 = \frac{R}{1 + \tilde{\tau}_1} s_0$$

where $s_0 = y_0 - \tilde{T}(y_0) - c_0$ is the total level of savings when there are no distortions in the economy. We can interpret $\tilde{\tau}$ as a capital income tax. It distorts inter-temporal consumption decisions by changing

[1] The second part of these notes is based on lecture notes by Florian Scheuer.

the relative price of c_0 and c_1. It can be interpreted as a wedge on the optimal savings decision. Notice that whenever $\tilde{\tau}_1 = 0$ we have no distortion in the inter-temporal choice of the agent.

Suppose the agent has separable preferences in consumption and labor of the form:

$$U(c_0, c_1, y_0) = u(c_0) + u(c_1) - h(y_0/w)$$

We can rewrite the preferences as $U(g(c_0, c_1), y_0)$ so that the utility is weakly separable in $g(\cdot)$ and y_0. It follows that Atkinson-Stiglitz applies: if non-linear income taxation is available the government finds optimal to set a flat zero tax on c_0 and c_1.[2]

2 Infinite Horizon Model - Chamley (1986)

In this section we introduce a model where capital returns and wages are endogenous. We focus on linear capital and labor taxes in an infinite horizon economy. There is aggregate uncertainty in the economy and each period t a state s_t is realized so that the history of aggregate uncertainty is a sequence $s^t = (s_0, s_1, \ldots, s_t)$. Output is produced according to a constant return to scale production function:

$$F\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right) \tag{2}$$

the productive capital at time t is the stock that was chosen at time $t-1$. The firm solves the following profit maximization problem:

$$\max_{K,L} F\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right) - w\left(s^t\right) L\left(s^t\right) - r\left(s^t\right) K\left(s^t\right)$$

Competitive labor and capital markets imply that input prices are equal to their marginal product:

$$w\left(s^t\right) = F_L\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right)$$

$$r\left(s^t\right) = F_K\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right)$$

The economy is populated by a representative agent whose utility is:

$$\sum_{s^t} \beta^t \Pr\left(s^t\right) u\left(c\left(s^t\right), L\left(s^t\right)\right)$$

where $\Pr\left(s^t\right)$ is the probability of history s^t. The aggregate resource constraint of the economy is:

$$c\left(s^t\right) + g\left(s^t\right) + K\left(s^t\right) - (1-\delta) K\left(s^{t-1}\right) \leq F\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right) \tag{3}$$

The output produced is employed to finance consumption, public spending and investments. The resource constraint implicitly assumes that aggregate uncertainty results from technology or government spending shocks. The government optimally chooses taxes on labor income $\tau^l\left(s^t\right)$ and taxes on capital income $\tau^k\left(s^t\right)$ and starts with initial debt B_0.

[2]Here is a short proof of the Atkinson Stiglitz result provided in Kaplow (2006). Suppose all individuals have weakly separable preferences $V(g(\mathbf{x}), y)$, where $\mathbf{x}$ is a vector of commodities. Suppose we start from a situation where there are positive taxes on commodities and we implement a policy such that $\mathbf{t} \to \mathbf{0}$: zero flat tax on all commodities. Suppose the government offsets the utility change of the agent with non-linear income taxes such that labor supply is unchanged at the optimum and $V(g(\mathbf{x}'), y) = V(g(\mathbf{x}), y)$. By definition every agent has the same utility as before and no one is willing to imitate another individual (if they were not willing to imitate before the tax change). By revealed preference we know that $\sum_k p_k x_k > y - T^*(y)$: the agent cannot afford the old bundle under the current income taxation. Under the old policy scenario the agent could afford the bundle and we had $\sum_k (p_k + \tau_k) x_k^0 \leq y - T(y)$. Combining the two inequalities we find that $T^*(y) > \sum_k \tau_k x_k^0 + T(y)$ and the total revenue raised after the tax change is strictly higher than total revenues before the tax change. Since incentive compatibility holds and we have no welfare effect by construction, the new policy is welfare improving since it raises more revenues.

We assume complete markets where the price of an Arrow-Debreu security is $p\left(s^t\right)$.[3] The government budget constraint is:

$$\sum_{t,s^t} p\left(s^t\right)\left[g\left(s^t\right)-\tau^l\left(s^t\right)w\left(s^t\right)L\left(s^t\right)-\tau^k\left(s^t\right)\left(r\left(s^t\right)-\delta\right)K\left(s^{t-1}\right)\right]\leq -B_0$$

Taxes on consumption and capital are employed to finance government layouts $g\left(s^t\right)$. Notice that the capital tax is levied on the capital gain net of the capital depreciation.

The household budget constraint reads:

$$\sum_{t,s^t} p\left(s^t\right)\left[c\left(s^t\right)+K\left(s^t\right)-w\left(s^t\right)\left(1-\tau^l\left(s^t\right)\right)L\left(s^t\right)-R\left(s^t\right)K\left(s^{t-1}\right)\right]\leq B_0 \tag{4}$$

where $R\left(s^t\right)=1+\left(1-\tau^k\left(s^t\right)\right)\left(r\left(s^t\right)-\delta\right)$ is the gross interest rate net of taxes. We can set up the Lagrangian for the consumer problem:

$$\begin{aligned} L &= \sum_{s^t}\beta^t \Pr\left(s^t\right)u\left(c\left(s^t\right),L\left(s^t\right)\right)+ \\ &\quad +\lambda\left[B_0-\sum_{t,s^t}p\left(s^t\right)\left[c\left(s^t\right)+K\left(s^t\right)-w\left(s^t\right)\left(1-\tau^l\left(s^t\right)\right)L\left(s^t\right)-R\left(s^t\right)K\left(s^{t-1}\right)\right]\right]\end{aligned}$$

The first order conditions are:

$$\beta^t \Pr\left(s^t\right)u_c\left(c\left(s^t\right),L\left(s^t\right)\right)-\lambda p\left(s^t\right)=0 \tag{5}$$

$$\beta^t \Pr\left(s^t\right)u_L\left(c\left(s^t\right),L\left(s^t\right)\right)+\lambda p\left(s^t\right)\left(1-\tau^l\left(s^t\right)\right)w\left(s^t\right)=0 \tag{6}$$

$$-\lambda p\left(s^t\right)+\lambda\sum_{s_{t+1}}p\left(s^{t+1}\right)R\left(s^{t+1}\right)=0 \tag{7}$$

On top of the FOCs, a non-arbitrage condition must hold between capital and Arrow-Debreu securities:

$$p\left(s^t\right)=\sum_{s_{t+1}}p\left(s^{t+1}\right)R\left(s^{t+1}\right) \tag{8}$$

We can define a competitive equilibrium as follows:

Definition: *A competitive equilibrium is a policy $\left\{g\left(s^t\right),\tau^k\left(s^t\right),\tau^l\left(s^t\right)\right\}$, an allocation $\left\{c\left(s^t\right),K\left(s^t\right),L\left(s^t\right)\right\}$ and prices $\left\{w\left(s^t\right),r\left(s^t\right),p\left(s^t\right)\right\}$, such that households maximizes utility s.t. budget constraint, firms maximize profits, the government budget constraint holds and markets clear*

Combining (5) and (6) we get the standard intratemporal condition for labor supply:

$$\frac{\beta^t \Pr\left(s^t\right)u_c\left(c\left(s^t\right),L\left(s^t\right)\right)}{p\left(s^t\right)}=-\frac{\beta^t \Pr\left(s^t\right)u_L\left(c\left(s^t\right),L\left(s^t\right)\right)}{p\left(s^t\right)\left(1-\tau^l\left(s^t\right)\right)w\left(s^t\right)}$$

$$\left(1-\tau^l\left(s^t\right)\right)w\left(s^t\right)=-\frac{u_L\left(c\left(s^t\right),L\left(s^t\right)\right)}{u_c\left(c\left(s^t\right),L\left(s^t\right)\right)} \tag{9}$$

[3] An Arrow-Debreu security is a financial instrument that provides one unit of consumption in a state s_t and zero units in any other state. We talk about complete markets whenever we can price such an asset in every state of the world.

From (5) and (7) we derive the so called Euler equation that pins down the slope of the consumption path of the agent:

$$u_c\left(c_0, L_0\right)=\frac{\beta^t \operatorname{Pr}\left(s^t\right) u_L\left(c\left(s^t\right), L\left(s^t\right)\right)}{p\left(s^t\right)} \tag{10}$$

Starting from (4), we can rewrite it using the optimality conditions and the no-arbitrage condition:

$$\begin{aligned}
\sum_{t, s^t} p\left(s^t\right)\left[c\left(s^t\right)-w\left(s^t\right)\left(1-\tau^l\left(s^t\right)\right) L\left(s^t\right)\right] &\leq B_0+R_0 K_0 \\
\sum_{t, s^t} \frac{\beta^t \operatorname{Pr}\left(s^t\right) u_c\left(c\left(s^t\right), L\left(s^t\right)\right)}{u_c\left(c_0, L_0\right)}\left[c\left(s^t\right)+\frac{u_L\left(c\left(s^t\right), L\left(s^t\right)\right)}{u_c\left(c\left(s^t\right), L\left(s^t\right)\right)} L\left(s^t\right)\right] &\leq B_0+R_0 K_0 \\
\sum_{t, s^t} \beta^t \operatorname{Pr}\left(s^t\right) u_c\left(c\left(s^t\right), L\left(s^t\right)\right)\left[c\left(s^t\right)+\frac{u_L\left(c\left(s^t\right), L\left(s^t\right)\right)}{u_c\left(c\left(s^t\right), L\left(s^t\right)\right)} L\left(s^t\right)\right] &\leq u_c\left(c_0, L_0\right)\left[B_0+R_0 K_0\right] \\
\sum_{t, s^t} \beta^t \operatorname{Pr}\left(s^t\right)\left[u_c\left(c\left(s^t\right), L\left(s^t\right)\right) c\left(s^t\right)+u_L\left(c\left(s^t\right), L\left(s^t\right)\right) L\left(s^t\right)\right] &\leq u_c\left(c_0, L_0\right)\left[B_0+R_0 K_0\right] \quad (11)
\end{aligned}$$

We call the constraint in (11) implementability constraint since it captures the agent's optimal choices subject to their feasibility.

Optimal Taxes The government chooses taxes to maximize the welfare of the representative agent subject to the resource constraint and the implementability constraint. The problem reads:

$$\max_{c\left(s^t\right), L\left(s^t\right), K\left(s^t\right), \tau_0^k} \sum_{s^t} \beta^t \operatorname{Pr}\left(s^t\right) u\left(c\left(s^t\right), L\left(s^t\right)\right)$$

s.t.

$$c\left(s^t\right)+g\left(s^t\right)+K\left(s^t\right)-(1-\delta) K\left(s^{t-1}\right) \leq F\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right)$$

$$\sum_{t, s^t} \beta^t \operatorname{Pr}\left(s^t\right)\left[u_c\left(c\left(s^t\right), L\left(s^t\right)\right) c\left(s^t\right)+u_L\left(c\left(s^t\right), L\left(s^t\right)\right) L\left(s^t\right)\right] \leq u_c\left(c_0, L_0\right)\left[B_0+R_0 K_0\right]$$

We assume τ_0^k is fixed, we denote with μ the multiplier on the implementability constraint and define:

$$W\left(c\left(s^t\right), L\left(s^t\right)\right)=u\left(c\left(s^t\right), L\left(s^t\right)\right)+\mu\left[u_c\left(c\left(s^t\right), L\left(s^t\right)\right) c\left(s^t\right)+u_L\left(c\left(s^t\right), L\left(s^t\right)\right) L\left(s^t\right)\right]$$

The government problem becomes:

$$\max_{c\left(s^t\right), L\left(s^t\right), K\left(s^t\right)} \sum_{s^t} \beta^t \operatorname{Pr}\left(s^t\right) W\left(c\left(s^t\right), L\left(s^t\right)\right)-\mu u_c\left(c_0, L_0\right)\left[B_0+R_0 K_0\right]$$

s.t.

$$c\left(s^t\right)+g\left(s^t\right)+K\left(s^t\right)-(1-\delta) K\left(s^{t-1}\right) \leq F\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right)$$

For any period $t \neq 0$ the FOCs are:

$$\beta^t \Pr\left(s^t\right) W_c\left(c\left(s^t\right), L\left(s^t\right)\right) - \gamma\left(s^t\right) = 0$$

$$\beta^t \Pr\left(s^t\right) W_L\left(c\left(s^t\right), L\left(s^t\right)\right) + \gamma\left(s^t\right) F_L\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right) = 0$$

$$-\gamma\left(s^t\right) + \sum_{s_{t+1}} \gamma\left(s^t, s_{t+1}\right)\left[F_k\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right) + (1-\delta)\right] = 0$$

Combining the FOCs we get an intratemporal condition:

$$-\frac{W_L\left(c\left(s^t\right), L\left(s^t\right)\right)}{W_c\left(c\left(s^t\right), L\left(s^t\right)\right)} = F_L\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right)$$

Using the household FOC in (6), we can rewrite the optimality condition as a function of the tax:

$$\tau^{l*}\left(s^t\right) = 1 - \frac{u_L\left(c\left(s^t\right), L\left(s^t\right)\right)}{W_L\left(c\left(s^t\right), L\left(s^t\right)\right)} \frac{W_c\left(c\left(s^t\right), L\left(s^t\right)\right)}{u_c\left(c\left(s^t\right), L\left(s^t\right)\right)} \tag{12}$$

From the government FOCs we can also derive an intertemporal condition:

$$W_c\left(c\left(s^t\right), L\left(s^t\right)\right) = \beta \sum_{s_{t+1}} \Pr\left(s^{t+1}|s^t\right) W_L\left(c\left(s^{t+1}\right), L\left(s^{t+1}\right)\right) R^*\left(s^{t+1}\right)$$

where $R^*\left(s^t\right) = 1 + F_K\left(K\left(s^{t-1}\right), L\left(s^t\right), s^t, t\right) - \delta$ is the untaxed gross return on capital net of depreciation. Again, exploiting the household's Euler equation we can derive:

$$R\left(s^{t+1}\right) = R^*\left(s^{t+1}\right) \frac{W_c\left(c\left(s^{t+1}\right), L\left(s^{t+1}\right)\right)}{u_c\left(c\left(s^{t+1}\right), L\left(s^{t+1}\right)\right)} \frac{u_c\left(c\left(s^t\right), L\left(s^t\right)\right)}{W_c\left(c\left(s^t\right), L\left(s^t\right)\right)} \tag{13}$$

Proposition 1: *Suppose that (i) there is no uncertainty (ii) there is a steady state. Then in the steady state $\tau^k = 0$ is optimal.*

It is easy to see that in a steady state when there is no uncertainty $R(ss) = R^*(ss)$, which implies:

$$1 + \left(1 - \tau^k(ss)\right)\left(F_k\left(K(ss), L(ss)\right) - \delta\right) = 1 + F_K\left(K(ss), L(ss), ss\right) - \delta$$

and that $\tau^k(ss) = 0$.

Now consider a special case with separable preferences and constant intertemporal elasticity of substitution:

$$u(c, L) = \frac{c^{1-\sigma}}{1-\sigma} - v(L)$$

then we have:

$$\begin{aligned} W(c, L) &= \frac{c^{1-\sigma}}{1-\sigma} - v(L) + \mu\left[c^{-\sigma}c - v'(L)L\right] \\ &= \left(\frac{1}{1-\sigma} + \mu\right)c^{1-\sigma} - \left[v(L) + \mu v'(L)L\right] \end{aligned}$$

therefore:

$$\begin{aligned} W_c &= (1+\mu(1-\sigma))\, c^{-\sigma} \\ &= (1+\mu(1-\sigma))\, u_c \end{aligned}$$

$$\frac{W_c}{u_c} = 1+\mu(1-\sigma)$$

Equation (13) reduces to $R\left(s^{t+1}\right) = R^*\left(s^{t+1}\right)$. Hence, we established that for separable preferences with constant intertemporal elasticity of substitution we have zero capital taxation even out of the steady state and in a model with uncertainty.

Tax Smoothing Take now the special case where $v(L) = \alpha L^{\gamma}/\gamma$ is isoelastic, we have:

$$W(c,L) = \left(\frac{1}{1-\sigma}+\mu\right) c^{1-\sigma} - \alpha\left(\frac{1}{\gamma}+\mu\right) L^{\gamma}$$

it follows that

$$\frac{W_L}{u_L} = 1+\mu\gamma$$

the optimal linear labor tax becomes:

$$\tau^{l*}\left(s^t\right) = 1 - \frac{1+\mu(1-\sigma)}{1+\mu\gamma}$$

Therefore, labor taxes are constant across states and over time. The government finds optimal to smooth distortions to labor supply. This result depends on the possibility of setting state-contingent capital taxes. If the labor elasticity is constant and shocks can be offset using capital taxes, there is no residual reason to differentially tax labor.

3 Infinite Horizon - Judd (1985)

We now introduce the model by Judd (1985) where the famous "zero steady state capital tax" result arise. We then show that the result is not general and depends on the agent's preferences as shown in Werning Straub (2015). Suppose there are two agents: capitalists and workers. The former save, get the return to capital and do not work; the latter supply one unit of labor inelastically and consume everything they earn. The government taxes return to capital and pays transfers to workers. Output is produced according to a constant return to scale technology with production function $F(k_t, n_t)$. Aggregate labor is $n_t = 1$ so that we can rewrite $f(k_t) = F(k_t, 1)$. Capitalists have utility $U(C_t)$ and workers' utility is $u(c_t)$. The resource constraint of the economy is:

$$c_t + C_t + g + k_{t+1} \leq f(k_t) + (1-\delta)\, k_t$$

Under the assumption of perfectly competitive labor markets, wages are:

$$w_t = F_L(k_t, n_t) = f(k_t) - k_t f'(k_t)$$

The after-tax return to capital is:

$$R_t = 1 + (1-\tau_t)(R_t^* - 1)$$

where $R^* = f'(k_t) + 1 - \delta$.

Capitalists solve the following maximization problem:

$$\max_{C_t, a_{t+1}} \sum_{t=0}^{+\infty} \beta^t U(C_t)$$

s.t.

$$C_t + a_{t+1} = R_t a_t$$

The optimality condition delivers the standard Euler equation $U'(C_t) = \beta R_{t+1} U'(C_{t+1})$. Since total wealth must equal total capital stock in equilibrium, using the Euler equation:

$$C_t + k_{t+1} = R_t k_t = \frac{U'(C_{t-1})}{\beta U'(C_t)} k_t$$

rearranging:

$$\beta U'(C_t)(C_t + k_{t+1}) = U'(C_{t-1}) k_t \tag{14}$$

Equation (14) is the implementability constraint. The government maximizes the following objective function, where γ is the Pareto weight on capitalists:

$$\max_{c_t, C_t, k_{t+1}} \sum_{t=0}^{\infty} \beta^t (u(c_t) + \gamma U(C_t))$$

The Lagrangian of the problem is:

$$\begin{aligned} L &= \sum_{t=0}^{\infty} \beta^t (u(c_t) + \gamma U(C_t)) \\ &\quad + \sum_{t=0}^{\infty} \beta^t \lambda_t (f(k_t) + (1-\delta) k_t - c_t - C_t - g - k_{t+1}) \\ &\quad + \sum_{t=0}^{\infty} \beta^t \mu_t (\beta U'(C_t)(C_t + k_{t+1}) - U'(C_{t-1}) k_t) \end{aligned}$$

with $\mu_0 = 0$ since there is no implementability in the first period (τ_0 is taken as given). The first order conditions wrt to c_t, k_{t+1} and C_t are respectively:

$$u'(c_t) = \lambda_t \tag{15}$$

$$\frac{\lambda_{t+1}}{\lambda_t}(f'(k_{t+1}) + 1 - \delta) = \frac{1}{\beta} + \frac{U'(C_t)}{\lambda_t}(\mu_{t+1} - \mu_t) \tag{16}$$

$$\mu_{t+1} = \frac{\mu_t}{k_{t+1}}\left(C_t + k_{t+1} + \frac{U'(C_t)}{U''(C_t)}\right) + \frac{1}{\beta k_{t+1}}\left(\gamma \frac{U'(C_t)}{U''(C_t)} - \frac{\lambda_t}{U''(C_t)}\right) \tag{17}$$

It is straightforward from equation (16) that whenever a steady state exists it involves zero capital taxes and $R(ss) = f'(k_t) + 1 - \delta = R^*$. This result is extremely powerful since it is independent of the welfare weight attached to capitalists. However, the result does not hold for the case where $\sigma = 1$. Rewrite the FOCs (16) and (17) using the inverse intertemporal elasticity of substitution $\sigma_t = -U''(C_t) C_t / U'(C_t)$ and defining the ratio $v_t = U'(C_t)/u'(c_t)$:

$$(16) \Rightarrow \frac{u'(c_{t+1})}{u'(c_t)}(f'(k_{t+1}) + 1 - \delta) = \frac{1}{\beta} + v_t(\mu_{t+1} - \mu_t)$$

$$(17) \Rightarrow \mu_{t+1} = \mu_t \left(\frac{1 - 1/\sigma}{k_{t+1}} + 1 \right) + \frac{1}{\beta k_t \sigma_t v_t} (1 - \gamma v_t)$$

Take the case where $\sigma = 1$ (log preferences) and the allocation converges to a steady state, then:

$$\mu_{t+1} - \mu_t = \frac{R^* - \frac{1}{\beta}}{v}$$

$$\mu_{t+1} - \mu_t = \frac{1 - \gamma v}{\beta k v}$$

$$\Rightarrow R^* - \frac{1}{\beta} = \frac{1 - \gamma v}{\beta k}$$

As long as there is a low enough weight on capitalists, capital is taxed in steady state. For a long time we thought that this was simply an anomaly for the logarithmic case. However, Werning and Straub (2015) show that the result does not hold for any $\sigma > 1$ by noticing that the steady state does not necessarily exist.

Proposition 2: *If $\sigma > 1$ and $\gamma = 0$, then for any initial k_0 the solution to the planning problem does not converge to the zero-tax steady state, or any other interior steady state.*

Suppose capital taxes are raised in the future, capitalists will decrease savings today for the substitution effect. A capital tax increase will also reduce agent's wealth and lower capitalists' consumption through the income effect. When $\sigma > 1$ the income effect prevails and capitalists save more. The increase in the capital stock increases wages and is beneficial for workers. For this reason the government wants to set positive capital taxes in the long-term. The opposite is true when $\sigma < 1$: the substitution effect is larger than the income effect and zero taxes in the future increase savings in the short term increasing wages and workers' consumption.

References

[1] Chamley, Christophe, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," Econometrica, 1986, 54 (3), pp. 607–622.

[2] Judd, Kenneth L., "Redistributive taxation in a simple perfect foresight model," Journal of Public Economics, 1985, 28 (1), 59 – 83.

[3] Kaplow, Louis, 2006. "On the undesirability of commodity taxation even when income taxation is not optimal," Journal of Public Economics, Elsevier, vol. 90(6-7), pages 1235-1250, August.

[4] Scheuer, Florian (2014), Lecture Notes (find at `http://web.stanford.edu/~scheuer/linear_capital_taxation.pdf`)

[5] Straub, Ludwig and Ivan Werning, "Positive Long Run Capital Taxation: Chamley-Judd Revisited," Working Paper, MIT 2015.