Economics 2450A: Public Economics
Section 10: Education Policies and Simpler Theory of Capital Taxation

Matteo Paradisi

November 14, 2016

In this section we study education policies in a simplified version of framework analyzed by Stantcheva (2016). We then review a simpler theory of capital taxation proposed by Saez and Stantcheva (2016) in a continuous time model.

1 Education Policies

We study a static model with human capital investments based on Bovenberg and Jacobs (2005, 2011) and Stantcheva (2016). Suppose individuals are heterogeneous in ability θ distributed according to $f(\theta)$. Agents can invest in education paying a monetary cost $M(e)$ such that $M'(e) > 0$ and $M''(e) \geq 0$. Wages are a function of ability and human capital that we can write as $w(\theta, e)$. Denote with $\rho_{\theta e}$ the Hicksian coefficient of complementarity between ability and education defined as:

$$\rho_{\theta e} = \frac{w_{\theta e} w}{w_\theta w_e}$$

Consider two special cases for the wage function: (i) multiplicative $w(\theta, e) = \theta e$ implies $\rho_{\theta e} = 1$; (ii) constant elasticity of substitution $w(\theta, e) = \left[\alpha \theta^{1-\rho} + (1-\alpha) e^{1-\rho}\right]^{\frac{1}{1-\rho}}$ implies $\rho_{\theta e} = \rho$. The elasticity of the wage to ability is $\varepsilon_{w,\theta} = \partial \log(w) / \partial \log(\theta)$.

Suppose utility is quasi-linear in consumption such that:

$$u(c, l) = c - h(l)$$

The agent consumes everything that is left after taxes and education investments such that $c(\theta) = w(\theta, e(\theta)) l(\theta) - M(e(\theta)) - T(w(\theta, e(\theta)) l(\theta), e(\theta))$. Solving the individual maximization problem we can define income and education wedges as:

$$\tau_y(\theta) = 1 - \frac{h'(l(\theta))}{w(\theta, e(\theta))} \tag{1}$$

$$\tau_e(\theta) = w_e(\theta, e(\theta)) l(\theta) (1 - \tau_y(\theta)) - M'(e(\theta)) \tag{2}$$

The elasticity of labor to the net of tax wage is:

$$\varepsilon = \frac{h'(l)}{h''(l) l}$$

We can write the indirect utility as $u(\theta) = c(\theta) - h(l(\theta))$. Using the Envelope we can derive the local incentive constraint:

$$\frac{\partial u(\theta)}{\partial \theta} = l(\theta)\frac{w_\theta(\theta, e(\theta))}{w(\theta, e(\theta))}h'(l(\theta)) \tag{3}$$

It differs from the one derived in the standard problem by the term $w_\theta(\theta, e(\theta))$, that takes into account the effect of ability on the wage. In the standard problem we normalized it to 1.

The resource constraint is:

$$\int (w(\theta, e(\theta))\, l(\theta) - u(\theta) - h(l(\theta)) - M(e(\theta)))\, f(\theta)\, d\theta \geq E \tag{4}$$

The government assigns welfare weight $\psi(\theta)$ to each θ and solves:

$$\max_{c(\theta), l(\theta), e(\theta), u(\theta)} \int \psi(\theta)\, u(\theta)\, f(\theta)\, d\theta$$

subject to (3) and (4).

The first order conditions wrt $u(\theta)$, $l(\theta)$ and $e(\theta)$ are respectively:

$$\psi(\theta) f(\theta) - \lambda f(\theta) = -\mu'(\theta) \tag{5}$$

$$\mu(\theta)\left[\frac{w_\theta}{w}h'(l) + l\frac{w_\theta}{w}h''(l)\right] + \lambda[w - h'(l)]\, f(\theta) = 0 \tag{6}$$

$$\mu(\theta)\left[-\frac{l}{w^2}w_e w_\theta h'(l) + \frac{l}{w}w_{\theta e}h'(l)\right] + \lambda[w_e l - M'(e)]\, f(\theta) \tag{7}$$

Taking the integral over (5) we derive $\mu(\theta) = \Psi(\theta) - F(\theta)$. Using (6), (1) and the definition of the elasticity of labor supply we get:

$$-\mu(\theta)\frac{w_\theta}{w}h'(l)\left[1 + l\frac{h''(l)}{h'(l)}\right] = \lambda w \tau_y(\theta) f(\theta)$$

$$-\mu(\theta) w_\theta (1 - \tau_y(\theta))\left[1 + l\frac{h''(l)}{h'(l)}\right] = \lambda w \tau_y(\theta) f(\theta)$$

$$\frac{\tau_y(\theta)}{1 - \tau_y(\theta)} = \frac{\Psi(\theta) - F(\theta)}{f(\theta)}\frac{w_\theta}{w}\left(\frac{1+\varepsilon}{\varepsilon}\right) = \frac{\Psi(\theta) - F(\theta)}{f(\theta)}\frac{\varepsilon_{w,\theta}}{\theta}\left(\frac{1+\varepsilon}{\varepsilon}\right) \tag{8}$$

The optimal income wedge is similar to the one we studied in Section 3. The formula has an extra term proportional to the elasticity of the wage to ability. Labor distortions are higher at the optimum when income is highly elastic to ability: the government distorts labor more when income is mostly explained by ability and less by effort or investment in education.

Notice that:

$$w_e l - M'(e) = \frac{\tau_e + M'(e)\tau_y}{1 - \tau_y}$$

Rearranging (7) and using (1) and (2) we get:

$$\mu(\theta)\, l\frac{w_e w_\theta}{w^2}h'(l)\left[-1 + \frac{w_{\theta e} w}{w_\theta w_e}\right] + \lambda\frac{\tau_e + M'(e)\tau_y}{1 - \tau_y}f(\theta) = 0$$

$$\frac{\tau_e(\theta) + M'(e)\tau_y(\theta)}{(1 - \tau_y(\theta))^2} = \frac{\Psi(\theta) - F(\theta)}{f(\theta)}\, l\frac{w_\theta w_e}{w}(1 - \rho_{\theta,e}) \tag{9}$$

The optimal education wedge decreases in the Hicksian coefficient of complementarity. When education and ability are complements (i.e. $\rho_{\theta e} > 0$) the government wants to discourage human capital investments in order to redistribute income more. On the other hand, if the coefficient is negative or low, education benefits low ability individuals more and a government subsidy to education helps redistribution. Suppose the wage function is $w(\theta, e) = \theta e$ and the monetary cost is linear in e, the optimal wedge is $\tau_e(\theta) = \tau_y(\theta)$. This is the special case studied by Bovenberg and Jacobs (2005) whose result is that income and education taxes are "Siamese Twins" and both margins should be distorted the same way. They also prove that the optimal linear education subsidy is equal to the optimal linear income tax rate, which is equivalent to making human capital expenses fully tax deductible.

2 A Simpler Theory of Capital Taxation

We introduce in this paragraph a continuous time model with wealth in the utility function. We study the case where utility is quasi-linear in consumption that allows us to transofrm the problem in a static taxation problem.

Suppose individual i has utility $u_i(c, k, z) = c + a_i(k) - h_i(z)$ where $a_i(\cdot)$ is increasing and concave and $h_i(\cdot)$ is the standard disutility from labor. Agents have heterogeneous discount rates δ_i. The discounted utility is:

$$V_i(\{c_i(t), k_i(t), z_i(t)\}) = \delta_i \int_0^{\infty} [c_i(t) + a_i(k_i(t)) - h_i(z(t))] e^{-\delta_i t} dt \tag{10}$$

Capital accumulates according to:

$$\frac{dk_i(t)}{dt} = rk_i(t) + z_i(t) - T(z_i(t), rk_i(t)) - c_i(t) \tag{11}$$

where $T(z_i(t), rk_i(t))$ is the tax paid by individual i and is dependent on income and capital returns. Wealth accumulation depends on the heterogeneous individual preferences, as embodied in the taste for wealth $a_i(\cdot)$ and in the impatience δ_i. It also depends on the net-of-tax return $\bar{r} = r(1 - T_k)$: capital taxes discourage wealth accumulation through a substitution effect (there are no income effects).

The Hamiltonian for the individual maximization problem is:

$$H_i(c_i(t), k_i(t), z_i(t), \lambda(t)) = [c_i(t) + a_i(k_i(t)) - h_i(z(t))] e^{-\delta_i t} + \lambda_i(t)[rk_i(t) + z_i(t) - T(z_i(t), rk_i(t)) - c_i(t)]$$

Taking first order conditions we have:

$$\frac{\partial H_i}{\partial c} = e^{-\delta_i t} - \lambda_i(t) = 0$$

$$\frac{\partial H_i}{\partial z_i} = -h_i'(z(t)) e^{-\delta_i t} + \lambda_i(t)[1 - T_z(z_i(t), rk_i(t))] = 0$$

$$\frac{\partial H_i}{\partial k_i(t)} = a_i'(k_i(t)) e^{-\delta_i t} + \lambda_i(t) r (1 - T_k(z_i(t), rk_i(t))) = -\lambda_i'(t)$$

Rearranging:

$$\lambda_i(t) = 1, \quad h_i'(z(t)) = 1 - T_z(z_i(t), rk_i(t)), \quad a_i'(k_i(t)) = \delta_i - r(1 - T_k(z_i(t), rk_i(t)))$$

Since utility is quasi-linear in consumption, the model converges immediately to a steady state. Denote (c_i, z_i, k_i) the steady state allocation, the problem collapses to a static optimization of the following objective function:

$$V_i\left(\{c_i(t), k_i(t), z_i(t)\}\right) = [c_i + a_i(k_i) - h_i(z)] + \delta_i\left(k_i^{init} - k_i\right)$$

where k_i^{init} is the inherited level of capital and $\left(k_i^{init} - k_i\right)$ is the utility cost of going from k_i^{init} to the steady-state level.

The government maximizes the following:

$$SWF = \int_i \omega_i U_i(c_i, k_i, z_i)\, di$$

with $\omega_i \geq 0$ is the Pareto weight on individual i. The social marginal welfare weight is $g_i = \omega_i U_{ic}$.

Optimal Linear Taxes Suppose the government sets linear income and capital taxes τ_L and τ_K. The individual chooses labor and capital according to $a_i'(k_i) = \delta_i - \bar{r}$ and $h_i'(z_i) = 1 - \tau_L$ with $\bar{r} = r(1-\tau_K)$. The government balances the budget through lump-sum transfers for a total of $G = \tau_K r k^m(\bar{r}) + \tau_L \cdot z^m(1-\tau_L)$, where $z^m(1-\tau_L) = \int_i z_i di$ is the aggregate labor income and $k^m(\bar{r}) = \int_i k_i$ is aggregate capital. Total consumption is $c_i = (1-\tau_K) r k_i + (1-\tau_L) z_i + \tau_K r k^m(\bar{r}) + \tau_L \cdot z^m(1-\tau_L)$ and the government maximizes:

$$SWF = \int_i \omega_i \left[(1-\tau_K) r k_i + (1-\tau_L) z_i + \tau_K r k^m + \tau_L \cdot z^m + a_i(k_i) - h_i(z_i) + \delta_i \cdot \left(k_i^{init} - k_i\right)\right] di$$

Using the Envelope theorem we get:

$$\begin{aligned}\frac{dSWF}{d\tau_K} &= \int \omega_i \left[-r k_i + r k^m + \tau_K r \frac{\partial k^m}{\partial \tau_k}\right] di \\ &= r k^m \left[\int_i \omega_i \left(1 - \frac{k_i}{k^m}\right) di - \frac{\tau_K}{1-\tau_K} e_K\right]\end{aligned}$$

where e_K is the elasticity of aggregate capital with respect to the net of tax return $\bar{r}$. At the optimum $dSWF/d\tau_K = 0$ and the optimal linear tax is:

$$\tau_k = \frac{1 - \bar{g}_K}{1 - \bar{g}_K + e_K}$$

where $\bar{g}_K = \int_i g_i k_i / \int_i k_i$. This is the standard formula for optimal linear taxes that we studied in Section 2 applied to capital. Notice that whenever capital accumulation is uncorrelated with social marginal welfare weights (i.e. $\bar{g}_K = 1$) the optimal tax is zero. The reason is that if capital has no tag value the government does not find optimal to tax capital for redistributive purposes. We also know from previous sections that the revenue maximizing tax rate corresponds to the case of $\bar{g}_K = 0$ and it is $\tau_K = 1/1 + e_K$.

Optimal Non-Linear Separable Taxes Suppose the government optimally sets $T_K(rk)$ and $T_L(z)$. The individual budget constraint is:

$$c_i = r k_i - T_K(r k_i) + z_i - T_L(z_i)$$

Define with $\bar{G}_K(rk)$ the average relative welfare weight on invividuals with capital income higher than rk. We have:

$$\bar{G}_K(rk) = \frac{\int_{\{i: r k_i \geq rk\}} g_i di}{P(r k_i \geq rk)}$$

Let $h_K(rk)$ be the distribution of capital income so that the Pareto parameter associated to the capital income distribution is:

$$\alpha_K(rk) = \frac{rk \cdot h_K(rk)}{1 - H_K(rk)}$$

Denote $e_K(rk)$ the elasticity of capital income with respect to the net of tax return $r(1 - T'_K(rk))$. Suppose the government introduces a small reform $\delta T_K(rk)$ where the marginal tax rate is increased by $\delta\tau_K$ in a small interval of capital income from rk to $rk + d(rk)$. The mechanical effect associated to the reform is:

$$d(rk)\,\delta\tau_K\,[1 - H_K(rk)]$$

The welfare effects just weights the mechanical effect by $\bar{G}(rk)$, the social marginal welfare weight associated to capital incomes above rk. Individuals who face the increase in the tax rate change their capital incomes by $\delta(rk) = -e_K\delta\tau_K/(1 - T'_K(rk))$. There are $h_K(rk)\,d(rk)$ individuals in the window affected by the tax change. Therefore, the total behavioral effect is:

$$-h_K(rk)\,d(rk)\,rk\frac{T'_K(rk)}{1 - T'_K(rk)}e_K(rk)$$

Summing up the three effects and rearranging we find:

$$\frac{T'_K(rk)}{1 - T'_K(rk)} = \frac{1}{e_K(rk)} \cdot \frac{1 - H_K(rk)}{rk \cdot h_K(rk)} \cdot \left(1 - \bar{G}_K(rk)\right)$$

Using the definition of the Pareto parameter we derive:

$$T'_K(rk) = \frac{1 - \bar{G}_K(rk)}{1 - \bar{G}_K(rk) + \alpha_K(rk) \cdot e_K(rk)}$$

which looks like the standard optimal non-linear tax formula.

References

[1] Bovenberg, L. and B. Jacobs. 2005. ìRedistribution and Education Subsidies are Siamese Twins,îJournal of Public Economics, 89(11-12), 2005-2035

[2] Bovenberg, L. and B. Jacobs. 2011. "Optimal Taxation of Human Capital and the Earnings Function," Journal of Public Economic Theory, 13, (6), 957-971.

[3] Saez, Emmanuel, and Stefanie Stantcheva. Working Paper. "A Simpler Theory of Optimal Capital Taxation." NBER Working Paper 22664.

[4] Stantcheva, Stefanie. "Optimal Taxation and Human Capital Policies over the Life Cycle," Journal of Political Economy, forthcoming.