Ramsey Taxation in OLG Economies

Jonathan Heathcote

January 24, 2012

1 Optimal Taxation in Over-lapping Generations Economies: Erosa and Gervais (JET, 2002)

Erosa and Gervais ask: What if households are not infinitely lived? In this case the households who might benefit from zero capital income tax in the long run are not the same households whose assets are confiscated initially.

Erosa and Gervais use the Ramsey approach to characterizing optimal taxation.

1.1 Ingredients of model

- Over-lapping generations
- All households within a cohort identical
- Individuals live $J+1$ periods from age 0 to J
- Population grows at rate n
- Share of age-j individuals in population, μ_j is time invariant
- Preferences defined over $c_{t,j}$ and $l_{t,j}$ where t indexes period of birth and j indexes age
- $j_0(t)$ is the age of the individual born at period t at date 0
- Date t aggregate labor input (per capita) is given by

$$l_t = \sum_{j=0}^{J} \mu_j z_j l_{t-j,,j}$$

- Output is CRS

$$y_t = f(k_t, l_t)$$

- Feasibility

$$c_t + (1+n)k_{t+1} - (1-\delta)k_t + g_t \leq y_t$$

- Policy instruments: proportional taxes on consumption, capital and labor income, and government debt
- Two alternatives are considered:
  1. Tax rates $\tau^a_{t-j,j}$, $\tau^w_{t-j,j}$, $\tau^c_{t-j,j}$ may vary with both time and age.
  2. Tax rates may vary only with time

Let

$$q_{t,j} = 1 + \tau^c_{t-j,j}$$
$$w_{t,j} = \left(1 - \tau^w_{t-j,j}\right) f_2(k_t, l_t)$$
$$r_{t,j} = \left(1 - \tau^a_{t-j,j}\right)\left(f_1(k_t, l_t) - \delta\right)$$

1.2 The individual's problem

An individual born at date $t \geq -J$ solves

$$\max U(c_{t,j_{0(t)}}, ..., c_{t,J}, l_{t,j_{0(t)}}, ...l_{t,J})$$

subject to

$$q_{t,j}c_{t,j} + a_{t,j+1} \leq w_{t,j}z_{t,j}l_{t,j} + (1 + r_{t,j})a_{t,j} \qquad j = j_0(t)...J$$

$$a_{t,j_0(t)} \text{ given and equal to } 0 \text{ if } t \geq 0$$

Assume that the utility function is time separable, and that the household discount factor is β.

Let U^t denote the indirect utility function of generation t.

The government's objective is

$$\max \sum_{t=-J}^{\infty} \gamma^t U^t$$

where γ is the rate at which the government discounts utility of future generations.

Definition 1 (Implementable Allocation)

Given $\{g_t\}_{t=0}^{\infty}$, $\{k_0,\, b_0\}$ and $\{a_{-j,j}\}_{j=1}^{J}$ where $k_0 + b_0 = \sum_{j=1}^{J} \mu_j a_{-j,j}$, an allocation $\left\{\{c_{t,j}, l_{t,j}\}_{j=j_0(t)}^{J},\ k_{t+J+1}\right\}_{t=-J}^{\infty}$ is implementable if there exists a fiscal policy $\left\{\{q_{t,j}, r_{t,j}, w_{t,j}\}_{j=j_0(t)}^{J},\ b_{t+J+1}\right\}_{t=-J}^{\infty}$ and a sequence of asset holdings $\left\{\{a_{t,j}\}_{j=j_0(t)}^{J}\right\}_{t=-J}^{\infty}$ such that:

1. Given prices from the fiscal policy, household decisions solve their problems
2. The government budget constraint is satisfied
3. Aggregate feasibility is satisfied

1.3 Proposition 1:

One of the fiscal policy instruments is redundant.

Thus without any loss of generality, it is possible to set $\tau^c_{t,j}$ equal to 0 for all t and j.

This result holds with or without age-dependent taxation

Note: P2.1a indicates that this equivalence relies on being able to adjust the capital tax rate in the first period. This is discussed on p348. Recall that in the infinite horizon set-up, the capital tax rate in the first period was set exogenously to prevent massive effectively lump-sum taxation at date zero. In the OG setup there is no need to restrict the initial capital tax rate: the government will not want to tax initial capital too heavily, because a front-loading taxation policy disproportionately hurts older cohorts at date 0.

The solution to the household problem

Let $p_{t,j}$ denote the Lagrange multiplier associated with the budget constraint faced by an age j individual born in t.

The necessary and sufficient conditions for a solution to the consumer's problem are

$$U_{t,c_{t,j}} - p_{t,j} = 0$$

$$U_{t,l_{t,j}} + p_{t,j} w_{t,j} z_j \leq 0$$

$$-p_{t,j} + p_{t,j+1}(1 + r_{t,j+1}) = 0$$

$$a_{t,J+1} = 0$$

1.4 Proposition 2:

An allocation $\left\{ \{c_{t,j}, l_{t,j}\}_{j=j_0(t)}^{J}, k_{t+J-1} \right\}_{t=-J}^{\infty}$ is implementable if and only if it satisfies feasibility and the implementability constraint, which is

$$\sum_{j=j_0(t)}^{J} \left(U_{c_{t,j}} c_{t,j} + U_{l_{t,j}} l_{t,j} \right) = U_{c_{t,j_0(t)}} \left(1 + r_{t,j_0(t)} \right) a_{t,j_0(t)}$$

for the cohort born in period t.

Proof:

(1) Implementable allocations satisfy feasibility and the implementability constraint

Multiply the household budget constraint by $p_{t,j}$ and sum over j

$$\sum_{j=j_0(t)}^{J} (p_{t,j} c_{t,j} - p_{t,j} w_{t,j} z_{t,j} l_{t,j}) = \sum_{j=j_0(t)}^{J} (p_{t,j}(1 + r_{t,j}) a_{t,j} - p_{t,j} a_{t,j+1})$$

Use the FOCs from the household problem to substitute out for prices.

(2) Allocations satisfying feasibility and implementability are implementable.

Suppose $\left\{\{c_{t,j}, l_{t,j}\}_{j=j_0(t)}^{J}, k_{t+J+1}\right\}_{t=-J}^{\infty}$ satisfies the feasibility and implementability constraints

Define

$$w_{t,j} = -\frac{U_{l_{t,j}}}{z_j U_{c_{t,j}}}$$

$$(1 + r_{t,j+1}) = \frac{U_{c_{t,j}}}{U_{c_{t,j+1}}}$$

$$p_{t,j} = U_{c_{t,j}}$$

By construction $\{c_{t,j}, l_{t,j}\}_{j=j_0(t)}^{J}$ satisfies the consumer's FOCs.

Want show that $a_{t,J+1} = 0$ (transversality condition) satisfied.

Given $a_{t,j_0(t)}$ iterate

$$a_{t,j+1} = w_{t,j} z_j l_{t,j} + (1 + r_{t,j}) a_{t,j} - c_{t,j}$$

forward until we have $a_{t,J+1}$. Then substitute in the expressions for factor prices into the RHS. The RHS can now be rearranged to give a version of the implementability constraint, implying that $a_{t,J+1} = 0$.

Note that further constraints must be added to the Ramsey problem if taxes cannot be made age-dependent.

1.5 The Ramsey problem

We are now finally in a position to state the Ramsey problem which is

$$\max_{\left\{\{c_{t,j}, l_{t,j}\}_{j=j_0(t)}^{J}, k_{t+J+1}\right\}_{t=-J}^{\infty}} \sum_{t=-J}^{\infty} \gamma^t W_t$$

subject to

$$c_t + (1+n)k_{t+1} - (1-\delta)k_t + g_t \leq y_t \qquad t = 0, ...$$

where

$$\begin{aligned} W_t \quad = \quad & \sum_{j=j_0(t)}^{J} \left[U_{t,,j} + \lambda_t \left(U_{c_{t,j}} c_{t,j} + U_{l_{t,j}} l_{t,j}\right)\right] - \\ & \lambda_t U_{c_{t,j_0(t)}} \left(1 + r_{t,j_0(t)}\right) a_{t,j_0(t)} \end{aligned}$$

γ is the intergenerational discount factor the government uses to weight the welfare of different generations

$\gamma^t \lambda_t$ is the multiplier associated with generation t's implementability constraint.

Let $\gamma^t \phi_t$ be the multiplier associated with the time t feasibility constraint.

1.6 Characterization of Optimal Fiscal Policies

The necessary conditions for a solution to the Ramsey problem are (wrt k_{t+1}, $c_{t,j}$ and $l_{t,j}$)

$$-\phi_t(1+n)+\gamma\phi_{t+1}(1-\delta+f_{k_{t+1}})=0 \tag{1}$$

$$\gamma^t W_{c_{t,j}}-\gamma^{t+j}\phi_{t+j}\mu_j=0 \tag{2}$$

$$\gamma^t W_{l_{t,j}}+\gamma^{t+j}\phi_{t+j}\mu_j z_j f_{l_{t+1}}=0$$

$$W_{l_{t,j}}+z_j f_{l_{t+j}} W_{c_{t,j}}=0 \tag{3}$$

What are the $W_{c_{t,j}}$ and $W_{l_{t,j}}$ terms? Assuming time-separable preferences (note: I would have expected additional terms for $j=j_0(t)$.)

$$\begin{aligned} W_{c_{t,j}} &= (1+\lambda_t)U_{c_{t,j}}+\lambda_t\left(U_{c_{t,j},c_{t,j}}c_{t,j}+U_{l_{t,j},c_{t,j}}l_{t,j}\right) \\ W_{l_{t,j}} &= (1+\lambda_t)U_{l_{t,j}}+\lambda_t\left(U_{c_{t,j},l_{t,j}}c_{t,j}+U_{l_{t,j},l_{t,j}}l_{t,j}\right) \end{aligned}$$

Thus eq. 3 can be written

$$-\frac{(1+\lambda_t)U_{l_{t,j}}+\lambda_t\left(U_{c_{t,j},l_{t,j}}c_{t,j}+U_{l_{t,j},l_{t,j}}l_{t,j}\right)}{(1+\lambda_t)U_{c_{t,j}}+\lambda_t\left(U_{c_{t,j},c_{t,j}}c_{t,j}+U_{l_{t,j},c_{t,j}}l_{t,j}\right)}=z_j f_{l_{t+j}}$$

and eq. 1 can be written

$$\begin{aligned} \frac{\gamma^t}{\gamma^{t+j}\mu_j}W_{c_{t,j}}(1+n) &= \gamma\frac{\gamma^t}{\gamma^{t+j+1}\mu_{j+1}}W_{c_{t,j+1}}(1-\delta+f_{k_{t+1}}) \\ \frac{\mu_{j+1}}{\mu_j}W_{c_{t,j}}(1+n) &= W_{c_{t,j+1}}(1-\delta+f_{k_{t+1}}) \\ W_{c_{t,j}} &= W_{c_{t,j+1}}(1-\delta+f_{k_{t+1}}) \end{aligned}$$

$$\frac{(1+\lambda_{t+1})U_{c_{t,j+1}}+\lambda_{t+1}\left(U_{c_{t,j+1},c_{t,j+1}}c_{t,j+1}+U_{l_{t,j+1},c_{t,j+1}}l_{t,j+1}\right)}{(1+\lambda_t)U_{c_{t,j}}+\lambda_t\left(U_{c_{t,j},c_{t,j}}c_{t,j}+U_{l_{t,j},c_{t,j}}l_{t,j}\right)}=\frac{1}{(1-\delta+f_{k_{t+1}})}$$

Compare these to the household's FOCs

$$-\frac{U_{l_{t,j}}}{U_{c_{t,j}}}=\left(1-\tau^w_{t-j,j}\right)f_2(k_t,l_t)z_j$$

$$\frac{U_{c_{t,j+1}}}{U_{c_{t,j}}}=\left(1+\left(1-\tau^a_{t-j,j}\right)\left(f_1(k_t,l_t)-\delta\right)\right)$$

It is clear that to get zero capital taxes we need

$$U_{c_{t,j},c_{t,j}}c_{t,j}+U_{l_{t,j},c_{t,j}}l_{t,j}\propto U_{c_{t,j}}$$

In general this condition will not be satisfied, even in a steady state, and thus the government will in general use non-zero taxes on both capital and labor income. It will not be satisfied because in general hours will vary with age – even in steady state – and thus the terms involving $l_{t,j}$ will vary with age.

This condition will be satisfied if preferences are separable between consumption and hours, and if utility is homothetic in consumption. This result parallels the infinite horizon example. In contrast to the infinite horizon example, however, with non-separable preferences we won't necessarily get zero capital taxes in steady state, because the planner is not trying to equate the marginal utility of consumption across age groups. Thus we have:

1.7 Proposition 3.3

With separable preferences that are homothetic in consumption, taxes on capital will be zero from time period 1 *and onwards.* **Note:** this result does not depend on the value for γ.

The next result is much weaker

1.8 Proposition 3.2

If Ramsey allocations converge to a steady state then in that steady state the tax on capital is zero if $\gamma = \beta(1+n)$ *and* $z_j = z$ *for all* j *(under either age-dependent or age-independent tax systems).*

The logic for this result isn't crystal clear (to me) in the paper. Here is my attempt. If $\gamma = \beta(1+n)$ then the steady state version of the planner's FOC for capital is

$$1 = \beta(1 - \delta + f_k)$$

This implies that if the planner sets a zero tax rate on capital, the household will choose constant consumption over the life-cycle. If, in addition, $z_j = z$ and the tax rate on labor is constant, chosen hours will also be independent of age. It follows that the Ramsey planner's inter-temporal FOC will be satisfied under this sort of tax scheme.

Erosa and Gervais then state some additional results:

- If the solution to the Ramsey problem converges to a steady state, the steady state has the modified golden rule property - from 1:

  $$1 - \delta + f_k = \frac{1+n}{\gamma}$$

  In other words, the steady state has the same capital-labor ratio that would be achieved if the social planner had access to lump-sum taxation.

- The steady state allocation is independent of the transition path leading to it

- To solve for the steady note that the steady state versions of 2 and 3 plus 1 and the steady state versions of the implementability constraint and the feasibility constraint constitute $2 \times (J+1) + 3$ equations and the same number of unknowns: $\left(\left\{c^j, l^j\right\}_{j=0}^{J}, \ k, \ \phi, \ \lambda\right)$

1.9 Proposition 4:

With additively separable preferences ($U(c, 1-l) = u(c) + v(l)$), if the Ramsey allocations converge to a steady state then in that steady state the relative tax rates on labor income at different ages are inversely related to the relative income elasticities of labor supplied at those ages, which in turn depends on the productivity profile.

1.10 Age independent taxes

To solve the problem when taxes cannot vary by age, we need to introduce some extra constraints for the Ramsey problem. In particular an allocation can only be implemented with age-independent taxation if: (1) the MRS between consumption and leisure is constant across individuals of different ages, and (2) the MRS between present and future consumption is constant across individuals of different ages. These requirements are imposed as additional constraints, and the Ramsey problem is resolved.

E & G find that:

- The set of allocations the government can implement with age-independent taxes is a proper subset of the set of implementable allcations under an age dependent tax system.
- If the solution to the Ramsey problem converges to a steady state, the steady state capital-labor ratio has the golden rule property and the steady state allocation is independent of the transition path.
- Even with additively separable preferences, capital income taxes are non-zero throughout transition and are only zero in steady state if $\gamma = \beta(1+n)$ and $z_j = z$ for all j. An intuition for this steady state result is that due to varying elasticity of labor supply, the government would like to tax labor at different rates at different ages. If it cannot do this directly it can achieve a similar outcome by using non-zero capital income taxes to effectively tax labor at different rates at different ages.

1.11 Numerical Examples (for the eventual steady state with utility separable in consumption and leisure)

- In terms of individual welfare or the behavior of aggregate variables, it makes little difference whether the tax system is age-dependent or age-independent.
- With age-dependent taxes, labor taxes vary significantly over the life cycle.
- Capital taxes are small (2-3 percent) but positive when taxes are age-independent.
- The equilibrium is highly sensitive to γ, the value of the intergenerational discount factor.

1.12 Example Based on Conesa, Kitao and Krueger

No population growth, households live for two periods.

Productivity 1 at age 1 and ε at age 2

Linear storage technology with return r

Labor produces output according to

$$L_t = l_{1,t} + \varepsilon l_{2,t}$$

This is interesting, because it clarifies that the issue is really about taxing savings, and not about taxing capital.

Household optimality conditions (with age-dependent taxes)

$$\begin{aligned}
\frac{U_{l_{1,t}}}{U_{c_{1,t}}} &= -(1-\tau_{l1,t}) \\
\frac{U_{l_{2,t+1}}}{U_{c_{2,t+1}}} &= -(1-\tau_{l2,t+1})\varepsilon \\
\frac{U_{c_{1,t}}}{U_{c_{2,t+1}}} &= \beta(1+r(1-\tau_{k,t+1}) = \frac{U_{l_{1,t}}}{U_{l_{2,t+1}}}\frac{(1-\tau_{l2,t+1})\varepsilon}{(1-\tau_{l1,t})}
\end{aligned}$$

Resource constraint

$$c_{1,t} + c_{2,t} + K_{t+1} - (1-\delta)K_t + G = rK_t + L_t$$

Let $\gamma^t \mu_t$ denote the multiplier on the resource constraint at t, and let $\gamma^t \lambda_t$ denote the multiplier on the implementability constraint for the generation born at t.

Suppose utility is of the form

$$\frac{c^{1-\sigma_1}}{1-\sigma_1} + \chi\frac{(1-l)^{1-\sigma_2}}{1-\sigma_2}$$

For the generation born at date t

$$\begin{aligned}
W_t \;=\; & \left[\frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} + \chi\frac{(1-l_{1,t})^{1-\sigma_2}}{1-\sigma_2} + \lambda_t\left(c_{1,t}^{1-\sigma_1} - \chi(1-l_{1,t})^{-\sigma_2} l_{1,t}\right)\right] \\
& + \beta\left[\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} + \chi\frac{(1-l_{2,t+1})^{1-\sigma_2}}{1-\sigma_2} + \lambda_t\left(c_{2,t+1}^{1-\sigma_1} - \chi(1-l_{2,t+1})^{-\sigma_2} l_{2,t+1}\right)\right]
\end{aligned}$$

The FOC wrt $c_{1,t}$ is

$$W_{c_{1,t}} = c_{1,t}^{-\sigma_1}(1+\gamma^t\lambda_t(1-\sigma_1)) = \gamma^t\mu_t$$

The FOC wrt $c_{2,t+1}$ is

$$W_{c_{2,t+1}} = \beta c_{2,t+1}^{-\sigma_1}(1+\gamma^t\lambda_t(1-\sigma_1)) = \gamma^{t+1}\mu_{t+1}$$

The FOC wrt K_{t+1} is

$$-\mu_t + \gamma\mu_{t+1}(1 - \delta + r) = 0$$

Combining these gives

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \beta(1 + r - \delta)$$

Note that this is true for any value for γ.

Thus in this case, the optimal capital income tax is zero in steady state, confirming one of the results in Erosa and Gervais

Now takes FOCs wrt labor supply to compute optimal labor tax rates:

$$W_{l_{1,t}} + W_{c_{1,t}} = 0$$

$$W_{l_{1,t}} = -\chi(1 - l_{1,t})^{-\sigma_2} - \lambda_t\chi(1 - l_{1,t})^{-\sigma_2} - \sigma_2\lambda_t\chi(1 - l_{1,t})^{-\sigma_2-1}l_{1,t}$$

So the intra-temporal FOC becomes

$$\chi(1 - l_{1,t})^{-\sigma_2}\left(1 + \lambda_t\left(1 + \sigma_2\frac{l_{1,t}}{(1 - l_{1,t})}\right)\right) = c_{1,t}^{-\sigma_1}(1 + \gamma^t\lambda_t(1 - \sigma_1))$$

and similarly for l_2 :

$$\chi(1 - l_{2,t+1})^{-\sigma_2}\left(1 + \lambda_t\left(1 + \sigma_2\frac{l_{2,t+1}}{(1 - l_{2,t+1})}\right)\right) = \beta c_{2,t+1}^{-\sigma_1}(1 + \gamma^t\lambda_t(1 - \sigma_1))\varepsilon$$

Taking ratios, we get

$$\frac{(1 - l_{1,t})^{-\sigma_2}\left(1 + \lambda_t\left(1 + \sigma_2\frac{l_{1,t}}{(1-l_{1,t})}\right)\right)}{(1 - l_{2,t+1})^{-\sigma_2}\left(1 + \lambda_t\left(1 + \sigma_2\frac{l_{2,t+1}}{(1-l_{2,t+1})}\right)\right)} = \frac{c_{1,t}^{-\sigma_1}}{\beta c_{2,t+1}^{-\sigma_1}}$$

The corresponding ratio from the households perspective is

$$\frac{(1 - l_{1,t})^{-\sigma_2}(1 - \tau_{l1,t})}{(1 - l_{2,t+1})^{-\sigma_2}(1 - \tau_{l2,t+1})} = \frac{c_{1,t}^{-\sigma_1}}{\beta c_{2,t+1}^{-\sigma_1}}$$

It follows that

$$\frac{1 - \tau_{l2,t+1}}{1 - \tau_{l1,t}} = \frac{1 + \lambda_t\left(1 + \frac{1}{\phi_1}\right)}{1 + \lambda_t\left(1 + \frac{1}{\phi_2}\right)}$$

where ϕ_i is the Frisch elasticity of labor supply at age $i : \phi_i = \frac{1-l_i}{l_i}\frac{1}{\sigma_2}$.

(to see that this is the Frisch elasticity, take

$$u_c w = u_l$$

take logs,

$$\ln u_c + \ln w = \ln \chi + \sigma_2 \ln(1-l)$$

and differentiate holding u_c constant)

So this expression says that labor taxes should fall with age iff $\phi_1 < \phi_2$, ie labor taxes should fall is the labor supply elasticity at age 2 is larger than at age 1. This in turn will be the case if hours worked are declining with age, since with these preferences, lower hours means a higher Frisch.

Now return to the agent's inter-temporal FOC:

$$\beta(1 + r(1 - \tau_{k,t+1})) = \frac{U_{l_{1,t}}}{U_{l_{2,t+1}}} \frac{(1 - \tau_{l2,t+1})\varepsilon}{(1 - \tau_{l1,t})}$$

From this it is immediate that a trend in labor taxes over the life cycle is equivalent in terms of the implied intertemporal wedge as a tax on capital income. We know that the optimum policy has a zero tax on capital, and - if hours decline over the life-cycle (as they will in Conesa et al.'s calibrated model) - labor taxes that decline with age. So in this constellation, the term $\frac{(1-\tau_{l2,t+1})}{(1-\tau_{l1,t})}$ is larger than one.

Suppose we were to legislate age-invariant labor tax rates. Then we would reduce the RHS of the agent's inter-temporal FOC. But we can generate the same inter-temporal wedge by reducing the term $(1-\tau_{k,t+1})$ on the LHS, which means taxing capital at a positive rate.