Capital Taxation

Laurence Jacquet

THEMA, OFS

September 2016

Outline

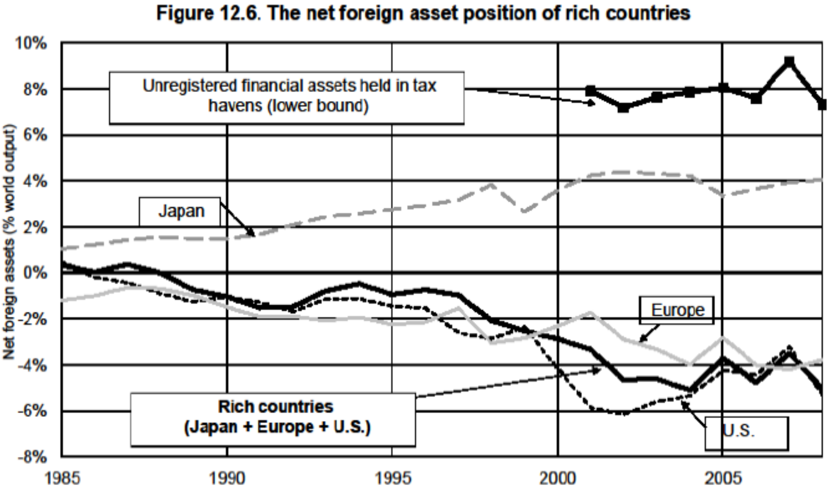
Optimal capital income taxation

Complex problem with many sub-literatures: Banks and Diamond in Mirrlees Review (2010) and Boadway (2012) provide recent surveys.

- Life-cycle models with linear and non-linear labor income taxes
- Models with future earnings uncertainty, e.g. New dynamic public finance (Kocherlakota, 2009)
- Many models with bequests or intergenerational taxation models [Skip]

In practice, it is very difficult to tax capital with capital mobility and little international coordination

- Without fiscal coordination (automated exchange of bank information, unified corporate tax base, etc.), all forms of capital taxation might well disappear in the long run, whatever the true social optimum might be.
- On these issues see e.g. Zucman (QJE 2013), Zucman (JEP 2014).



Unregistered financial assets heldin tax havens are higher than the official net foreign debt of rich countries.
Sources and series: see piketty.pse.ens.fr/capital21c.

Table 1

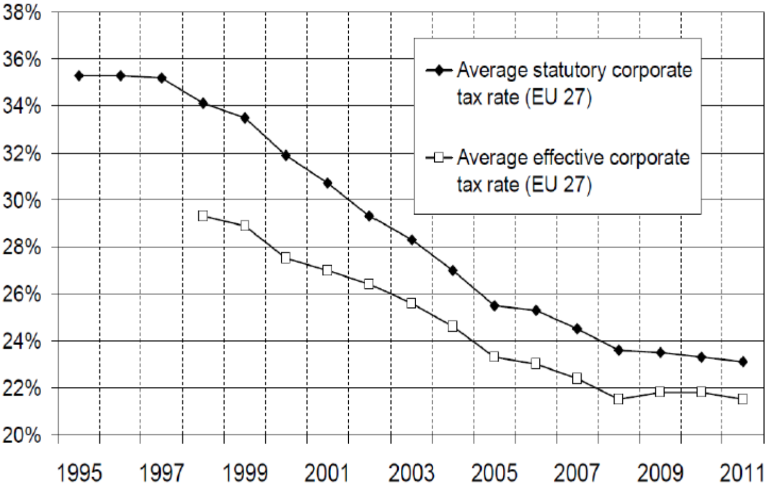
The World's Offshore Financial Wealth

	Offshore wealth ($ billions)	*Share of financial wealth held offshore*	*Tax revenue loss ($ billions)*
Europe	2,600	10%	75
United States	1,200	4%	36
Asia	1,300	4%	35
Latin America	700	22%	21
Africa	500	30%	15
Canada	300	9%	6
Russia	200	50%	1
Gulf countries	800	57%	0
Total	**7,600**	**8.0%**	**190**

Source: Author's computations (see Zucman 2013a, b) and online Appendix.
Notes: Offshore wealth includes financial assets only (equities, bonds, mutual fund shares, and bank deposits). Tax revenue losses only include the evasion of personal income taxes on investment income earned offshore as well as evasion of wealth, inheritance, and estate taxes.

Source: Piketty's PSE website

Corporate tax competition in the EU



Source: Taxation trends in the EU, Eurostat 2011

- From now on we assume closed economy (or perfect international coordination): not because this is realistic, but because in order to know whether we should coordinate, we need to know what would be the coordinated optimum (some people believe that even if perfect coordination was possible, we should have zero capital tax for purely economic reasons).
- In effect, capital mobility & limited coordination raise elasticities of capital supply; see e.g. discussion of income-shifting/tax-avoidance elasticity in Piketty, Saez, Stantcheva (AEJ 2014).
- In the case of perfect mobility with zero international coordination, this elasticity $= \infty \Rightarrow$ No capital tax is possible.

- In practice, there are always frictions and asset specificities (e.g. some capital equipment cannot move easily and/or is more valuable in certain territories), so this elasticity $< \infty$; but it can get quite high, and could keep rising in the future.

Outline

Chamley-Judd zero capital tax rate

Chamley Judd: Problem
Chamley (Ecta 1986) and Judd (JPubE 1985) considered a dynamic Ramsey problem with:

- Identical agents;
- Endogenous capital accumulation;
- An exogenous stream of government expenditures that must be financed from linear taxes on labor income and capital income
- What is the least distortionary way to finance the government expenditures?
- They solved the problem under full commitment. The government solves the optimal policy problem at time 1 and commits not to change its future policy.

Life-cycle model and zero capital tax rate

Life-cycle model with:

1. Finite horizon
2. Infinite horizon

Life-cycle model with finite horizon

Suppose:

- Preferences are $u(x_1, \ell_1) + \beta u(x_2, \ell_2)$ with β : discount factor
- Wage rate is identical in both periods
- $\beta = \frac{1}{1+r}$ (steady state)

$\Rightarrow$ Optimal: $\tau_K = 0$, $t_1 = t_2$ and $t_{w_1} = t_{w_2}$
$\Rightarrow$ $x_1 = x_2$ and $\ell_1 = \ell_2$ (steady state)

Optimal for capital taxes to be zero in the long run in a representative agent dynamic model.

Proof of zero capital tax rate with finite horizon

Boadway (2013, p.90).
Assume discount factor $\beta = \frac{1}{1+r}$ and also assume wage rate identical in both periods.
All prices and taxes are in present value terms.
Consumer prices for goods and leisure are: $q_1 = 1$, $q_2 = p_2 + t_2$, $\omega_1 = w_1 + t_{w_1}$, $\omega_2 = w_2 + t_{w_2}$ where producer prices are $p_1 = 1$, p_2, w_1, w_2.

Individual maximization problem:

$$\underset{x_1, x_2, \ell_1, \ell_2}{Max} \; u(x_1, \ell_1) + \beta u(x_2, \ell_2)$$
$$\text{s.to: } x_1 + q_2 x_2 = \omega_1 \ell_1 + \omega_2 \ell_2 \text{ (mult. } \alpha)$$

FOCs (x_1, x_2):

$$u_x^1 = \alpha \text{ and } \beta u_x^2 = \alpha q_2$$

FOCs (ℓ_1, ℓ_2):

$$u_\ell^1 = -\alpha \omega_1 \text{ and } \beta u_\ell^2 = -\alpha \omega_2$$

From these, the individual's budget constraint can be rewritten as

$$x_1 \frac{u_x^1}{\alpha} + \frac{\beta u_x^2}{\alpha} x_2 = -\ell_1 \frac{u_\ell^1}{\alpha} - \frac{\beta u_\ell^2}{\alpha} \ell_2$$
$$\Leftrightarrow x_1 u_x^1 + \beta u_x^2 x_2 = -u_\ell^1 \ell_1 - \beta u_\ell^2 \ell_2$$

Lagrangian for the government:

$$
\begin{aligned}
& L\left(x_1, x_2, \ell_1, \ell_2, \lambda, \alpha\right) \equiv u\left(x_1, \ell_1\right) + \beta u\left(x_2, \ell_2\right) \\
& +\lambda\left[w_1 \ell_1 + w_2 \ell_2 - x_1 - p_2 x_2 - R\right] \\
& +\gamma\left[x_1 u_x^1 + \beta u_x^2 x_2 + u_\ell^1 \ell_1 + \beta u_\ell^2 \ell_2\right]
\end{aligned}
$$

FOCs (x_1, x_2, ℓ_1, ℓ_2):

$$u_x^1 - \lambda + \gamma\left[u_x^1 + u_{xx}^1 x_1 + u_{\ell x}^1 \ell_1\right] = 0 \qquad (x_1)$$

$$\beta u_x^2 - \lambda p_2 + \gamma\beta\left[u_x^2 + u_{xx}^2 x_2 + u_{\ell x}^2 \ell_2\right] = 0 \qquad (x_2)$$

$$u_\ell^1 + \lambda w_1 + \gamma\left[u_\ell^1 + u_{x\ell}^1 x_1 + u_{\ell\ell}^1 \ell_1\right] = 0 \qquad (\ell_1)$$

$$\beta u_\ell^2 + \lambda w_2 + \gamma\beta\left[u_\ell^2 + u_{x\ell}^2 x_2 + u_{\ell\ell}^2 \ell_2\right] = 0 \qquad (\ell_2)$$

Proof of zero capital tax rate with finite horizon (cont'd)

By the stated assumptions: $p_2 = \beta \left(= \frac{1}{1+r}\right)$ and $w_2 = \beta w_1$, conditions (x_2) and (ℓ_2) become:

$$u_x^2 - \lambda + \gamma \left[u_x^2 + u_{xx}^2 x_2 + u_{\ell x}^2 \ell_2\right] = 0 \qquad (x_2')$$
$$u_\ell^2 + \lambda w_2 + \gamma\beta \left[u_\ell^2 + u_{x\ell}^2 x_2 + u_{\ell\ell}^2 \ell_2\right] = 0 \qquad (\ell_2')$$

$\Rightarrow$ (x_1), (x_2'), (ℓ_1), (ℓ_2') satisfied if $x_1 = x_2$ and $\ell_1 = \ell_2$ (i.e. both consumption and labor supply are constant over time, and there is no saving). So $u_x^1 = u_x^2$ and $u_\ell^1 = u_\ell^2$.

Using individual FOCs:

$$\frac{u_x^2}{u_x^1} = \frac{q_2}{\beta} = 1 = \frac{p_2}{\beta}$$

and

$$\frac{u_\ell^2}{u_\ell^1} = \frac{\omega_2}{\beta\omega_1} = 1 = \frac{w_2}{\beta w_1}$$

$\Rightarrow q_2 = p_2$ **i.e. no tax on capital income.**

$\Rightarrow \frac{q_2}{q_1} = \frac{w_2}{w_1}$ **i.e. labor taxes are same over time.**(As well $c_1 = c_2$ and $\ell_1 = \ell_2$)

Life-cycle model with infinite horizon (Chamley-Judd)

An infinitely lived representative individual/household faces the following problem:

$$\underset{\{x_t,\ell_t\}_{t=1}^{\infty}}{Max} \sum_{t=1}^{\infty} \beta^{t-1} u\left(C_t, L_t\right)$$

$$\text{s.to:} \qquad A_{t+1} = \left(1 + \left(1 - \tau_t^K\right) r_t\right) A_t + \left(1 - \tau_t^L\right) W_t L_t - C_t$$

where C_t denotes consumption, L_t labor supply, β the discount factor, A_t the risk-free asset holding, W_t the wage rate, r_t and the interest rate. τ_t^K The government uses:
τ_t^K a **linear** tax on the income from capital received at t and
τ_t^L a **linear** tax on the income from labor received at t.

Note: The intertemporal utility function does not impose weak separability between consumption and labor.

Chamley-Judd: Individual

The individual problem can be solved **recursively:**

$$V(A_t) = \max_{A_{t+1},L_t} u\left(\underbrace{\left(1+\left(1-\tau_t^K\right)r_t\right)A_t + \left(1-\tau_t^L\right)W_t L_t - A_{t+1}}_{C_t}, L\right.$$

$$+$$

The envelope condition is

$$V'(A_t) = \left(1+\left(1-\tau_t^K\right)r_t\right)u_C(C_t, L_t)$$

The FOC w.r.to A_{t+1} is:

$$u_C(C_t, L_t) = \beta V'(A_{t+1})$$

Combining these two equations yields the **Euler equation**:

$$u_C(C_t, L_t) = \beta\left(1+\left(1-\tau_{t+1}^K\right)r_{t+1}\right)u_C(C_{t+1}, L_{t+1})$$

Moreover, the FOC w.r.to L_t is:

$$\left(1-\tau_t^L\right) W_t u_C\left(C_t, L_t\right) = -u_L\left(C_t, L_t\right)$$

Chamley-Judd: Firm

There is a representative firm that produces output from capital and labor under a constant returns to scale technology $F(K_t, L_t)$.
The firm faces the following profit maximization problem:

$$\max_{K_t, Lt} F(K_t, L_t) - R_t K_t - W_t L_t$$

where R_t is the rental rate of capital, which is equal to the interest rate r_t plus the depreciation rate δ, i.e. $R_t \equiv r_t + \delta$.

At the optimum, factor prices are equal to their marginal product and the firm makes zero profits:

$$\begin{aligned} R_t &= F_K(K_t, L_t) \\ W_t &= F_L(K_t, L_t) \end{aligned}$$

Let me scrutinize the Euler equation:

$$u_C\left(C_t, L_t\right) = \beta\left(1 + \left(1 - \tau^K_{t+1}\right) r_{t+1}\right) u_C\left(C_{t+1}, L_{t+1}\right)$$

Assume $\tau^K_{t+1} = 0$:

$$\frac{u_C\left(C_t, L_t\right)}{u_C\left(C_{t+1}, L_{t+1}\right)} = \beta\left(1 + r_{t+1}\right) = \frac{1 + r_{t+1}}{1 + \rho}$$

If the **interest rate** r_{t+1} is **above the rate of time preference or discount rate** ρ with $\beta = 1/(1+\rho)$, then agents choose to **accumulate capital** and to **postpone their consumption indefinitely** $(C_t < C_{t+1} < C_{t+2} < ...)$ and this cannot be a steady-state.

Conversely, if the interest rate is **below** the rate of time preference ρ, agents choose to **desacumulate capital, i.e. to borrow, indefinitely** and to consume more today ($C_t > C_{t+1} > C_{t+2} > ...$). This cannot be a steady-state either.

$\Rightarrow$ The long-run steady-state interest rate r is equal to the time preference/discount rate ρ.

From $R_t = F_K(K_t, L_t)$, we also know that at the steady-state capital must be such that $F_K(K, L) - \delta = r$ (since $R_t \equiv r_t + \delta$).

Chamley-Judd: Government

The government needs to finance an exogenous stream of government expenditures $\{G_t\}_{t=1}^{\infty}$ from linear taxes on capital and labor $\{\tau_t^K, \tau_t^L\}_{t=1}^{\infty}$. The indebtedness is equal to B_t at time t and it evolves according to:

$$B_{t+1} = (1 + r_t) B_t + G_t - \tau_t^K r_t A_t - \tau_t^L W_t L_t$$

where G_t are (exogenous) gvt expenditures at time t.
The assets A_t held by the indivdiuals are either capital K_t or government debt B_t , i.e. $A_t = K_t + B_t$.

Chamley-Judd: Government (cont'd)

Substituting the government budget constraint into the individual budget constraint yields the resource constraint (by Walras' law):

$$\begin{aligned} K_{t+1} &= (1+r_t)\,K_t + W_t L_t - C_t - G_t \\ &= (1-\delta)\,K_t + R_t K_t + W_t L_t - C_t - G_t \\ &= (1-\delta)\,K_t + F(K_t, L_t) - C_t - G_t \end{aligned}$$

where the 2nd line was obtained using $(1+r_t) = 1-\delta+R_t$ from $R_t = r_t + \delta$ and the last line was obtained by using the zero profit condition.

Chamley-Judd: Competitive equilibrium

Definition of a **competitive equilibrium**:
It is a feasible allocation $\{K_t, L_t, C_t\}_{t=1}^{\infty}$, a price system $\{W_t, r_t\}$ and a government policy $\{\tau_t^K, \tau_t^L\}_{t=1}^{\infty}$ such that:

- Given the price system and the government policy, the allocation solves both the firm's problem and the individual's problem.
- The allocation satisfies the resource constraint.

Note: For **most government policies** $\{\tau_t^K, \tau_t^L\}_{t=1}^{\infty}$, there is **no competitive equilibrium** since the tax levels are not appropriate to finance the government expenditures $\{G_t\}_{t=1}^{\infty}$.

Chamley-Judd: Competitive equilibrium

The set of competitive equilibria, which can be indexed by the government policy $\left\{\tau_t^K, \tau_t^L\right\}_{t=1}^{\infty}$, is fully determined by:

- The two optimality conditions of the firm's problem
- The two optimality conditions of the individual's problem
- The household's budget constraint
- The resource constraint.

Chamley-Judd: Primal approach

The **Ramsey taxation problem** consists in selecting the competitive equilibrium which maximizes the utility of the indvidual.

To solve the Ramsey problem, we shall rely on the **primal approach** which proceeds as follows:

1. Determine the restrictions imposed by linear taxes on the set of allocations that can be sustained as a competitive equilibrium. This yields a **set** of **implementable** allocations.
2. **Select** the implementable allocation that **maximizes social welfare**.
3. **Recover** the **optimal tax rates** implied by the optimal allocation.

Chamley-Judd: Primal approach

The intertemporal budget constraint of the individual is:

$$\left(1+\left(1-\tau_1^K\right)r_1\right)[K_1+B_1]+\sum_{t=1}^{\infty}\frac{\left(1-\tau_t^L\right)W_tL_t-C_t}{\prod_{i=2}^{t}\left(1+\left(1-\tau_i^K\right)r_i\right)}=0$$

The **restrictions imposed by linear taxes are captured by the FOCs of the individual and of the firm**:

$$1+\left(1-\tau_t^K\right)r_t=\frac{u_C\left(C_{t-1},L_{t-1}\right)}{\beta u_C\left(C_t,L_t\right)},\quad \left(1-\tau_t^L\right)W_t=-\frac{u_L\left(C_t,L_t\right)}{u_C\left(C_t,L_t\right)}$$

and $r_t=F_K\left(K_t,L_t\right)-\delta$

Chamley-Judd: Primal approach

Substituting these conditions into the individual's intertemporal budget constraint yields:

$$\left[u_C\left(C_1, L_1\right)\left[1+\left(1-\tau_1^K\right)\left(F_K\left(K_1, L_1\right)-\delta\right)\right]\right]\left[K_1+B_1\right]$$
$$+\sum_{t=1}^{\infty} \beta^{t-1}\left[u_L\left(C_t, L_t\right) L_t-u_C\left(C_t, L_t\right) C_t\right]=0$$

This equation is known as the **implementability constraint**.

Chamley-Judd: Primal approach

Together, the **implementability constraint and the resource constraint are necessary and sufficient** to characterize the set of allocations which can be sustained as a **competitive equilibrium**.

Proof:

- Necessity: By construction, a competitive equilibrium must satisfy the implementability constraint.

Chamley-Judd: Primal approach

Proof (cont'd):

- Sufficiency: We need to show that an allocation $\{K_t^*, L_t^*, C_t^*\}_{t=1}^{\infty}$ that satisfies the two constraints can be decentralized as a competitive equilibrium.
  - Define $r_t = F_K(K_t^*, L_t^*) - \delta$, $W_t = F_L(K_t^*, L_t^*)$, τ_K^t and τ_L^t such that the individual's optimality conditions hold.
  - Using these equations, it is straightforward to recover the individual's budget constraint from the implementability constraint. Thus, $\{K_t^*, L_t^*, C_t^*\}_{t=1}^{\infty}$ does satisfy the individual's budget constraint.
  - Thus, the allocation $\{K_t^*, L_t^*, C_t^*\}_{t=1}^{\infty}$ is profit maximizing for the firm and welfare maximizing for the individual. Moreover, it satisfies the resource constraint.

Chamley-Judd: Primal approach

The **optimal allocation** of resources can be obtained by **maximizing welfare subject to the implementability constraint and to the resource constraint**.

The primal approach therefore yields the following Lagrangian:

$$L \equiv \sum_{t=1}^{\infty} \beta^{t-1} \left\{ u\left(C_t, L_t\right) + \Psi\left[u_L\left(C_t, L_t\right) L_t - u_C\left(C_t, L_t\right) C_t\right] \right\}$$

$$+\theta_t \left[\left(1-\delta\right) K_t + F(K_t, L_t) - C_t - G_t - K_{t+1}\right]$$

$$+\Psi \left[u_C\left(C_1, L_1\right) \left[1 + \left(1 - \tau_1^K\right)\left(F_K\left(K_1, L_1\right) - \delta\right)\right]\right] \left[K_1 + B_1\right]$$

with Ψ as mult. for the implementability constr. and θ_t as mult. for the resource constraint.

Chamley-Judd: Primal approach

The FOCs w.r.to C_t and K_{t+1} are, respectively

$$u_C\left(C_t, L_t\right) + \Psi\left[u_L\left(C_t, L_t\right) L_t - u_{CC}\left(C_t, L_t\right) C_t - u_C\left(C_t, L_t\right)\right] = \theta_t$$

$$\beta\theta_{t+1}\left[1 - \delta + F_K(K_{t+1}, L_{t+1})\right] = \theta_t$$

Chamley-Judd: Optimal policy

We want to derive the optimal policy in **steady state**. This requires constant government expenditures over time, i.e. $G_t = G$ for all t.
In **steady-state**, $C_t = C$, $L_t = L$ and $K_t = K \Rightarrow$

- Hence, the FOC w.r.to Ct implies that $\theta_t = \theta$.

- The **FOC w.r.to K_{t+1}** simplifies to:

$$\beta[1-\delta+F_K(K,L)]=1 \text{ or } \frac{1+r}{1+\rho}=1 \text{ (since } F_K(K,L)-\delta=r)$$

In steady state, the **Euler equation** can be written as:

$$\beta\left(1+\left(1-\tau_K\right)r\right)=1 \text{ or } \frac{\left(1+\left(1-\tau_K\right)r\right)}{1+\rho}=1$$

or, equivalently, as

$$\beta\left(1+\left(1-\tau_K\right)\left(F_K(K,L)-\delta\right)\right)=1$$

Hence, since $\beta[1-\delta+F_K(K,L)]=1$ (FOC w.r.to K_{t+1}) optimality requires

$$\tau_K=0,$$

i.e. **no taxation of capital income (in the long run)!**

Chamley-Judd: Intuition

Chamley (1986) and Judd (1985) theorem: In the steady state of an **infinite** horizon general equilibrium model, government expenditures should be exclusively financed from taxes on labor income.

Intuition:

- In steady state, the net return on capital $(1-\tau_K)\,r$ is always equal to the discount rate (or rate of time preference) ρ where $\beta = 1/(1+\rho)$.
- Thus, the **capital tax is entirely shifted to workers**:
  - The capital stock falls such that capitalists keep earning the same rate of return.
  - The gross wage rate of workers falls.

Chamley-Judd: Intuition (cont'd)

- This is $\simeq$ to having an **infinite long-run elasticity of capital supply**: any infinitesimal change in the net interest rate generates a saving response that is unstainable in the long run, unless the net interest rate returns to its initial level $r = \rho$. (This zero capital tax result breaks down whenever the **long run elasticity of labor supply is finite**.)
- A small (constant) tax on capital cumulates to ∞ over a long enough horizon; the price of consumption at t relative to consumption at $t + T$ is multiplied by $\left(\frac{1+r}{1+r(1-\tau_K)}\right)^T$.

Chamley-Judd: Criticisms

- Agents do not have infinite lives.

⇒ **Reply to this criticism:** It is possible to have a **dynastic interpretation of the infinite horizon**. Ricardian equivalence notion of an overlapping-generation model of representative individuals in each period who make operative bequests to their immediate heirs based on altruism. In such a model, the path of consumption over all cohorts can be replicated by maximizing a dynastic utility function subject to an intertemporal budget constraint (Barro JPE 1974):

-
  - A household consists of a dynasty where each generation lives for one period.
  - Parents are altruistic and attach a weight β to the welfare of their children (who themselves care about the welfare of their own children).
  - **If bequests are altruistic motivated, then the Chamley-Judd result implies that they should not be taxed**.

Chamley-Judd: Criticisms (cont'd)

However, this **dynastic model depends on a number of far-fetched assumptions** (Boadway 2012, pp.90-92):

- The Ricardian model depends on being able to aggregate all individuals in any cohort into a representative agent whose bequest motive applies to the representative agent of the next cohort.

This is logically inconsistent; Bernheim and Bagwell (JPE 1988) show the implausibility of such an aggregation: Each child has two parents and parents come from different families $\Rightarrow$ The heirs of any given person of the current cohort are shared indirectly with all other persons of the same cohort (i.e. bequest motive becomes dissipated), so making a bequest amounts to making a voluntary contribution to a national public good ($\Rightarrow$ underprovision of bequest).

Chamley-Judd: Criticisms (cont'd)

- The utility discount factor of the representative agent must be the same as that of the intergenerational social planner. That is, the weight put on the utility of future generations must be dictated by the extent of altruism alone. Not obvious why it should be the case.

Time consistent linear taxation

This literature solves the problem under full commitment. The government solves the optimal policy problem at time 1 and commits not to change its future policy.

- Taxpayers take long-run and short-run decisions
- Long-run decisions, like saving, create asset income that is fixed in the future
- Short-run decisions, like labor supply, create income in the same period

Time consistent linear taxation (cont'd)

- Second-best optimal tax policy is determined before long-run decisions are taken
- **Second-best tax policies are generally time-inconsistent**: even benevolent governments will choose to change tax policies after long-run decisions are undertaken

If households anticipate such re-optimizing, the outcome will be inferior to the second-best
Governments may implement policies up front to mitigate that problem.

Time consistent linear taxation (cont'd): An illustrative model

Based on Fischer (RevEconDyn&Control 1981), and Persson and Tabellini (survey in Handbook of Public Economics 2002).

- Two periods, two goods (x_1, x_2) and labor in period 2 (ℓ).
- Quasilinear utility: $u(x_1) + x_2 + h(1-\ell)$, $u(.)$, $h(.)$ strictly concave
- Time endowment: 1, wealth endowment: 1
- Wage rate $=$ 1, interest rate $=$ 0
- Second-period taxes: τ_k, τ_ℓ on k, ℓ.
- Fixed government revenue G

Time consistent linear taxation (cont'd): An illustrative model

Consumer problem:

$$\underset{\{x_1,\ell\}}{Max}\ u(x_1) + (1 - t_\ell)_\ell + (1 - t_k)(1 - x_1) + h(1 - \ell)$$

$$\Rightarrow x_1\left(t_k\right),\ x_1'\left(t_k\right) > 0,\ k\left(t_k\right) = 1 - x_1\left(t_k\right);\ \ell\left(t_\ell\right),\ \ell'\left(t_\ell\right) < 0.$$

Indirect utility: $v(t_k, t_\ell \ell)$, with $v_{t_k} = -(1 - x_1)$, $v_{t_\ell} = -\ell$.

Time consistent linear taxation (cont'd): An illustrative model

Government Policy:

$\underset{\{t_k, t_\ell\}}{Max}\ v(t_k, t_\ell)$ s.to: $t_\ell \ell(t_\ell) + t_k k(t_k) = G$

Second-best tax:

$$\frac{t_\ell}{1-t_\ell} = \frac{\kappa}{\eta_\ell(t_\ell)} > 0, \ \frac{t_k}{1-t_k} = \frac{\kappa}{\eta_k(t_k)} > 0$$

where $\eta_\ell(t_\ell) = \frac{\partial \ell}{\partial t_\ell}\frac{t_\ell}{\ell}$, $\eta_k(t_k) = \frac{\partial k}{\partial t_k}\frac{t_k}{k}$ and κ is a constant.

Time consistent linear taxation (cont'd): An illustrative model

- Ex post, government will reoptimize by treating k as fixed and set t_k as high as possible (e.g. $t_k = 1$).
- Individuals anticipate this and reduce saving.
- Time-consistent equilibrium is inferior to second-best.
- Government may react by providing ex ante saving incentives.
- Inability to commit may be responsible for high capital income and wealth tax rates in practice.
- Widespread use of investment and savings incentives.
- Same phenomenon applies to human capital investment, investment by firms, and housing.

Outline

Life-cycle model: Atkinson-Stiglitz

Heterogeneous individuals
Government uses non-linear taxes on earnings, should the government also use taxes on savings?

Result: If utility is weakly separable and tastes are homogeneous, then the government should use only labor income tax and should not use tax on savings.

It is possible to generate a Pareto improvement by eliminating the taxation of capital.

This is Atkinson and Stiglitz theorem (JPubE 1976), Konishi (JPubE 1995), Laroque (EL 2005), Kaplow (JPubE 2006).

Life-cycle model: Atkinson-Stiglitz, basic two-period, two-skill case

See Boadway (2012, pp.96-100)

x_i^j : consumption in period j by type i $(i, j = 1, 2)$
$\ell_i^1 = y_i^1 / w_i$: labor supply by type i **in period** 1 **only**, $w_2 > w_1$
Utility: $u\left(x_i^1\right) - h\left(\ell_i^1\right) + \beta u\left(x_i^2\right)$

Government problem (full commitment assumed):

$$Max\ n_1 \left(u\left(x_1^1\right) - h\left(\frac{y_1^1}{w_1}\right) + \beta u\left(x_1^2\right) \right) + n_2 \left(u\left(x_2^1\right) - h\left(\frac{y_2^1}{w_2}\right) + \beta u\left(x_2^2\right) \right)$$

Life-cycle model: Atkinson-Stiglitz, basic two-period, two-skill case

s.to:

$$n_1 \left(y_1^1 - x_1^1 - \frac{x_1^2}{1+r} \right) + n_2 \left(y_2^1 - x_2^1 - \frac{x_2^2}{1+r} \right) = R \quad (\text{mult. } \lambda)$$

and an **IC on a lifetime basis** assume IC applies to type-2's):

$$u\left(x_2^1\right) - h\left(\frac{y_2^1}{w_2}\right) + \beta u\left(x_2^2\right) \geq u\left(x_1^1\right) - h\left(\frac{y_1^1}{w_2}\right) + \beta u\left(x_1^2\right) \quad (\text{mult. } \gamma)$$

Basic two-period, two-skill case

FOCs on consumption

$$
\begin{aligned}
(n_1 - \gamma)\, u'\left(x_1^1\right) - \lambda n_1 &= 0 \\
(n_1 - \gamma)\, \beta u'\left(x_1^2\right) - \frac{\lambda n_1}{1+r} &= 0 \\
(n_2 + \gamma)\, u'\left(x_2^1\right) - \lambda n_2 &= 0 \\
(n_2 + \gamma)\, \beta u'\left(x_2^2\right) - \frac{\lambda n_2}{1+r} &= 0
\end{aligned}
$$

Basic two-period, two-skill case

$\Rightarrow$ An intertemporal wedge expression:

$$\frac{u'\left(x_1^1\right)}{\beta u'\left(x_1^2\right)} = \frac{u'\left(x_2^1\right)}{\beta u'\left(x_2^2\right)} = \frac{1}{1+r}$$

i.e. the intertemporal MRS in consumption for both skill-types should equal the intertemporal MRT $(1+r)$.
$\Rightarrow$ **No tax on capital income or savings** (which is equivalent to a tax on 2nd-period consumption); Atkinson-Stiglitz theorem applies.

Intuition in line with the ICC (Boadway, 2012, p.97-98**)**
Denote c_1 the disposable income (recall that labor in period 1 only). The choice of (y_1, c_1), by an individual represents the choice of a point along the lifetime budget constraint, $c_1 = y_1 - T(y_1)$.
Individuals use their disposable income to finance first- and second-period consumption according to their budget constraint

$$c_1 = x_1 + \frac{x_2}{1+r}$$

Given the separable utility function, high-wage individuals will choose the same consumption profile as low-wage individuals if they have the same disposable income c_1.

$\Rightarrow$ High-skill mimicking low-skill will choose the same (x_1, x_2) bundle, and it is this that renders **differential taxation of present and future consumption useless as a policy instrument alongside the nonlinear income tax**.

Similar Atkinson-Stiglitz model but labor supply in both periods

- **Labor supply in both periods**
- $\Rightarrow$ Now the government chooses consumption and labor in both periods (j) (x_i^j, y_i^j) for $i, j = 1, 2$.
- The government can impose a **nonlinear labor income tax in each period.**
- The government can impose a **differential commodity tax on 1st- and 2nd-period consumption, or equivalently a capital income tax or subsidy**.

- Lifetime IC (assume IC applies to type-2's):

$$\begin{aligned} & u\left(x_1^2\right) - h\left(y_1^2/w^2\right) + \beta\left[u\left(x_2^2\right) - h\left(y_2^2/w^2\right)\right] \\ \geq \quad & u\left(x_1^1\right) - h\left(y_1^1/w^2\right) + \beta\left[u\left(x_2^1\right) - h\left(y_2^1/w^2\right)\right] \end{aligned}$$

$\Rightarrow$ From FOCs: **no intertemporal distortion on consumption for either skill type**. AS theorem applies.

Proof:

$$L \equiv \sum_{i=1,2} n_i \left(u\left(x_i^1\right) - h\left(\frac{y_i^1}{w_i}\right) + \beta u\left(x_i^2\right) - \beta h\left(\frac{y_i^2}{w_i}\right)\right)$$
$$+\lambda \left\{ \sum_{i=1,2} n_i \left(y_i^1 + \frac{y_i^2}{1+r} - x_i^1 - \frac{x_i^2}{1+r} - R \right) \right\}$$
$$+\gamma \left\{ u\left(x_2^1\right) - h\left(y_2^1/w^2\right) + \beta\left[u\left(x_2^2\right) - h\left(y_2^2/w^2\right)\right]\right.$$
$$\left. -u\left(x_1^1\right) + h\left(y_1^1/w^2\right) - \beta\left[u\left(x_1^2\right) - h\left(y_1^2/w^2\right)\right]\right\}$$

FOCs w.r.to x_2^1, x_2^2, x_1^1, x_1^2 (we skip the other FOCs)

$$\left(n_2+\gamma\right)u'\left(x_2^1\right)=\lambda n_2$$

$$\left(n_2+\gamma\right)\beta u'\left(x_2^2\right)=\frac{\lambda n_2}{1+r}$$

$$\left(n_1-\gamma\right)u'\left(x_1^1\right)=\lambda n_1$$

$$\left(n_1-\gamma\right)\beta u'\left(x_1^2\right)=\frac{\lambda n_1}{1+r}$$

$$\Rightarrow \frac{u'\left(x_1^1\right)}{\beta u'\left(x_1^2\right)}=\frac{u'\left(x_2^1\right)}{\beta u'\left(x_2^2\right)}=\frac{1}{1+r}$$

No intertemporal distortion on consumption (i.e. no differential taxation of 1st- and 2nd-period consumption) for either skill type.
$\Rightarrow$ **no tax on savings/capital**.

Critical assumptions behind this result of no tax on capital with nonlinear labor income taxation

- The government can commit (individuals reveal their types in 1st period!), for a very interesting discussion regarding this: see Boadway (2012) pp.98-99.
- Separable form of the utility function: intertemporally and intratemporally
- This result applies even **if wage rates for the two types vary over the two periods** (see later "uncertain future wage rates").
- This assumes that the government can impose **age-specific tax systems**!

Earnings in both periods with wages that vary; age-dependent taxation

Assume

- **Wages** in period j are w_i^j for $i, j = 1, 2$.
- **No uncertainty**
- Identical preferences: $u(x^1) - h(\ell^1) + \beta u(x^2) - \beta h(\ell^2)$
- Government can commit to two-period tax system
- Fully nonlinear tax on present and future income
- Assume incentive constraint applies to type-2's

Earnings in both periods with wages that vary; age-dependent taxation

The Government's problem is identical to the previous one.

$$L \equiv \sum_{i=1,2} n_i \left(u\left(x_i^1\right) - h\left(\frac{y_i^1}{w_i}\right) + \beta u\left(x_i^2\right) - \beta h\left(\frac{y_i^2}{w_i}\right)\right)$$
$$+\lambda \left\{ \sum_{i=1,2} n_i \left(y_i^1 + \frac{y_i^2}{1+r} - x_i^1 - \frac{x_i^2}{1+r} - R\right)\right\}$$
$$+\gamma \left\{ u\left(x_2^1\right) - h\left(y_2^1 / w^2\right) + \beta \left[u\left(x_2^2\right) - h\left(y_2^2 / w^2\right)\right]\right.$$
$$\left. -u\left(x_1^1\right) + h\left(y_1^1 / w^2\right) - \beta \left[u\left(x_1^2\right) - h\left(y_1^2 / w^2\right)\right]\right\}$$

Earnings in both periods with wages that vary; age-dependent taxation

$$\Rightarrow \frac{u'\left(x_1^1\right)}{\beta u'\left(x_1^2\right)} = \frac{u'\left(x_2^1\right)}{\beta u'\left(x_2^2\right)} = \frac{1}{1+r}$$

No intertemporal distortion on consumption (i.e. no differential taxation of 1st- and 2nd-period consumption) for either skill type.

$\Rightarrow$ **no tax on savings/capital**.

Tax Smoothing

From conditions on x_2^j, y_2^j:

$$\frac{h'\left(y_2^1 / w_2^1\right)}{u'\left(x_2^1\right) w_2^1} = 1 = \frac{h'\left(y_2^2 / w_2^2\right)}{u'\left(x_2^2\right) w_2^2}$$

$\Rightarrow$ **Tax smoothing** for (high-ability) type 2's. $T' = 0$ in the case of the high-ability persons, in both periods.

Of course, that does not imply that the average tax rate is the same in both periods: given that the wage rate differs over periods, the average tax rate certainly differs.

Tax Smoothing

For (low-ability) type 1's, $T' > 0$ and it varies between periods. An exception: if r**elative wages of the two types are the same in both periods, and the utility of labor supply is constant elasticity (Diamond 2007)**:

$$\frac{h'\left(y_1^1/w_1^1\right)w_1^2}{h'\left(y_1^2/w_1^2\right)w_1^1}\Delta = \beta\left(1+r\right),$$

$$\text{where } \Delta = \left(\frac{w_2^2}{w_2^1}\right)^{\sigma+1}\frac{n_1 w_2^{1\sigma+1} - \gamma w_1^{1\sigma+1}}{n_1 w_2^{2\sigma+1} - \gamma w_1^{2\sigma+1}}$$

$\Rightarrow$ Tax smoothing for $1'$s if $\Delta = 1$, i.e. if $\frac{w_1^2}{w_1^1} = \frac{w_2^2}{w_2^1}$ (identical age-earnings profiles, assuming $h'\left(\ell_i\right) = \ell_i^\sigma$).

If $\frac{w_1^2}{w_1^1} < \frac{w_2^2}{w_2^1}$, marginal tax rate for 1's rises over time.

Different discount rates/tastes for saving

Suppose $\beta_1 \neq \beta_2$, so government objective becomes

$$Max \; n_1 \left(u\left(x_1^1\right) - h\left(\frac{y_1^1}{w_1}\right) + \beta_1 u\left(x_1^2\right) \right) + n_2 \left(u\left(x_2^1\right) - h\left(\frac{y_2^1}{w_2}\right) + \beta_2 u\left(x_2^2\right.\right.$$

and the IC is

$$u\left(x_2^1\right) - h\left(\frac{y_2^1}{w_2}\right) + \beta_2 u\left(x_2^2\right) \geq u\left(x_1^1\right) - h\left(\frac{y_1^1}{w_2}\right) + \beta_2 u\left(x_1^2\right) \quad (\text{mult. } \gamma)$$

Different discount rates/tastes for saving (cont'd)

FOCs yield

$$\frac{u'\left(x_2^1\right)}{\beta_2 u'\left(x_2^2\right)} = \frac{1}{1+r} \lesseqgtr \frac{u'\left(x_1^1\right)}{\beta_1 u'\left(x_1^2\right)} \text{ if } \beta_1 \lesseqgtr \beta_2$$

Intertemporal decisions of the high-skilled (type 2) remain undistorted, while those of the low-skilled (type 1) are distorted if $\beta_1 \neq \beta_2$.

Different discount rates/tastes for saving (cont'd)

Saez (JPubE 2002) and Diamond and Spinnewijn (AEJ: EconPol 2011) argue that the evidence suggests a **positive correlation between wage rates and the weight individuals put on future utility**, so $\beta_2 > \beta_1$.

In this case where high-skilled types tend to have higher savings rates than the low-skilled, a positive capital income tax should be imposed on the low-skilled. That is, implicit tax on savings of low-wage types (Saez, JPubE 2002 and Diamond and Spinnewijn AEJ: EconPol 2011).

Intuition: Taxing savings of low-wage types reduces their second-period consumption, makes it more costly for high-wage types to mimic given their **lower** utility discounting.

With linear tax on savings (dual income tax), case for positive linear tax since high-ability have higher savings rates

Uncertain future wage rates, Cremer and Gahvari (EJ 1995)

- **Cremer and Gahvari (EJ 1995)** consider the Atkinson-Stiglitz theorem in an economy in which **wage rates are uncertain** and individuals must make some consumption decisions (for durable goods as housing) before they know their wage rates while other goods are purchased and labor supplied after wage rates are known.
- No need of a dynamic approach, the intuition of their result can be shown in a static model in which decisions are taken sequentially. See Boadway (2012, pp.75-77).
- Since some goods are purchased before uncertainty is resolved, the AS theorem fails.

Uncertain future wage rates, Cremer and Gahvari (EJ 1995)

- No differential taxation among goods purchased ex post and **goods purchased ex ante** bear a different and **lower tax rate** (Cremer and Gahvari EJ 1995).
- Intuition: inducing **all** persons to increase their consumption of the durable good makes it more difficult for those who turn out to have higher skills ex post to mimic those with low skills, **since their ex post consumption requirements are higher**.
- Cremer and Gahvari argue that this provides justification for the preferential treatment that is offered to housing and consumer durables in many income tax systems.
- This argument is closely related to those in favor of taxing capital income in a dynamic setting where there is wage uncertainty in e.g. Golosov, Tsyvinski and Werning (NBER Macro 2007).

Uncertain future wage rates

Assumptions:

- **Labor supply in both periods**
- Two periods, 1 and 2
- **Common wage w_1 in period 1,**
- and either w_1^2 (low wage) or w_2^2 (high wage) in period 2 (incentive constraint in period 2 only); no aggregate uncertainty, so **a given proportion of population turns out to be high-wage and the rest are low-wage**.

- n_i^2 = distribution of i 's in period 2
- Lifetime expected utility:

$$u(x^1) - h\left(y^1/w^1\right) + \beta \sum_{i=1,2} n_i^2 \left(u\left(x_i^2\right) - h\left(\frac{y_i^2}{w_i^2}\right)\right)$$

- There are nonlinear labor income taxes in both periods now!
- This case strains (again) the assumption of **commitment since individuals will reveal their types in the 1st period**!
- **Differential commodity tax on 1st- and 2nd- period consumption**, or **equivalently a capital income tax or subsidy**.

Summary of tax instruments:

- The government imposes a **common lump-sum income tax in the 1st period** since all persons are indistinguishable then.
- In the 2nd period, **a nonlinear labor income tax** can be imposed as well as **a tax on capital income** (based on first-period saving).

Remarks:

- Although wage uncertainty is in principle insurable, it is assumed that Insurance markets do not exist, possibly because of information constraints facing insurers.
- The government announces its tax structure at the beginning of the first period and again is assumed to be able to commit to it.

First-best economy:
The government provides full consumption insurance through the tax system.

Second-best economy

- Incentive constraint $\Rightarrow$ underinsurance ($x_2^2 > x_1^2$) in the 2nd period; high-wage persons must be given more consumption to preclude them from mimicking the low-skilled.
- It becomes optimal to **tax capital income**. Intuitively, reducing saving makes it more costly for the high-wage to mimic the 2nd-period labor income of the low-wage types by reducing the amount of 2nd-period consumption made available by saving.

Uncertain future wage rates (cont'd)

Government problem:

$$Max \quad u(x^1) - h\left(y^1/w^1\right) + \beta \sum_{i=1,2} n_i^2 \left(u\left(x_i^2\right) - h\left(\frac{y_i^2}{w_i^2}\right)\right)$$

s.to:

$$y^1 - x^1 + \frac{1}{1+r} \sum_i n_i^2 \left(y_i^2 - x_i^2\right) \geq R \quad (\text{mult. } \lambda)$$

and a second-period IC:

$$u\left(x_2^2\right) - h\left(\frac{y_2^2}{w_2^2}\right) \geq u\left(x_1^2\right) - h\left(\frac{y_1^2}{w_2^2}\right) \quad (\text{mult. } \gamma)$$

FOCs on x_1 and x_i^2:

$$u'(x^1) = \lambda \tag{1}$$

$$\beta n_1^2 u'(x_1^2) - \lambda \frac{1}{1+r} n_1^2 - \gamma u'\left(x_1^2\right) = 0 \tag{2}$$

$$\beta n_2^2 u'\left(x_2^2\right) - \lambda \frac{1}{1+r} n_2^2 + \gamma u'(x_2^2) = 0 \tag{3}$$

Summing the LHS of (2) and (3) as well as their RHS, we obtain:

$$(1+r)\left[\beta\left[n_1^2 u'\left(x_1^2\right) + n_1^2 u'\left(x_2^2\right)\right] - \gamma\left[u'\left(x_1^2\right) - u'\left(x_2^2\right)\right]\right] = \lambda$$

Therefore, using (1), we obtain:

$$\frac{u'\left(x^1\right)}{\beta\left(1+r\right)\sum_i n_i^2 u'\left(x_i^2\right)-\gamma\left(1+r\right)\left(u'\left(x_1^2\right)-u'\left(x_2^2\right)\right)}=1$$

From IC (that yields $-\left(u'\left(x_1^2\right) - u'\left(x_2^2\right)\right) < 0$) and FOCs:

$$\frac{u'\left(x^1\right)}{\beta\left(1+r\right)\sum_i n_i^2 u'\left(x_i^2\right)} < 1 \tag{4}$$

This says that there is a **positive intertemporal consumption wedge**, implying that there should be a positive tax on saving.

Reducing saving makes it harder for type 2 to mimic the income of type 1 in period 2 (by reducing the amount of consumption made available by saving).

This intuitive result carries forward to the case when individuals are **heterogeneous in the first period**, see Golosov, Kocherlakota and Tsyvinski (ReStud 2003).

Dynamic macro public finance literature

Workers are exposed to the risk of losing their skills (which are private info) and, hence, to face wage cuts.

- What is the **optimal provision of insurance against the skill risk**?

This is the main focus of the dynamic macro public finance literature (New Dynamic Public Finance).

There are two parts to the problem:

- What is the **optimal** incentive-feasible **allocation** of resources?
  - This is the best allocation that can be implemented by a direct-truthful mechanism.
- How could the optimal allocation be **implemented** by the government in a decentralized economy using realistic fiscal instruments (instead of a direct mechanism)?

Dynamic macro public finance literature

There is a continuum of mass 1 of agents.
For simplicity, agents are assumed to be **ex-ante identical** and, hence, we have a **pure social insurance problem**.
Allowing for ex-ante heterogeneity would add a redistribution dimension to the optimal policy problem (see before).
Each individual maximizes his expected lifetime utility:

$$\mathbf{E}\left[\sum_{t=1}^{T}\beta^{t-1}\left[u\left(x_t\right)-v\left(\ell_t\right)\right]\right]$$

where T is the life-span of the individual, β is his discount factor, x_t his consumption at t and ℓ_t his labor supply at t.
Note the change of notation: time is now in subscrib.

Assume that $u'(.) > 0$, $u''(.) < 0, v'(.) > 0$ and $v''(.) > 0$.
An individual of age t with productivity w_t produces output $y_t = w_t \ell_t$
y_t is observable while w_t and ℓ_t are private information.

Nature draws a skill vector $w^T = (w_1, w_2, ..., w_T)$ for each agent according to a probability measure μ_W

- w^T represents the agent's lifetime sequence of skills;
- The draws are identically and independently distributed across agents;
- At any time t, an agent only knows his history of skills up to time t and denoted by $w_t = (w_1, ..., w_t)$.

By the law of large numbers, the fraction of agents with history w^T is determined by μ_W.

There is an exogeneous risk-free interest rate r.

- It is possible to endogenize the accumulation of capital, in which case the interest rate is equal to the marginal product of capital net of depreciation.

Planner's Problem

By the revelation principle, the planner's problem is to find the best allocation implementable by a direct truthful mechanism.
The allocation given to a worker claiming to be of type w^T is:

$$\left\{x_t\left(w^T\right), y_t\left(w^T\right)\right\}_{t=1}^T$$

Under a direct mechanism, each worker must choose a reporting strategy $\sigma(.)$ which specifies the reported skills for each possible realization of his own skill vector.

- Under the reporting strategy $\sigma(.)$ a worker of type $\overline{w}^T$ obtains the allocation designed for workers of type $\sigma\left(\overline{w}^T\right)$.
- The skill σ_t reported at time t must be w^t measurable.

For any w^T , the reporting strategy can be written as
$\sigma\left(w^T\right) = \left(\sigma_1\left(w^1\right), \sigma_2\left(w^2\right), ..., \sigma_T\left(w^T\right)\right)$

- A truth-telling strategy $\sigma^*(.)$ is characterized by $\sigma^*(w^T) = w^T$ for all w^T.

Planner's Problem

Let $\sigma(.)$ denote the reported skills up to time t.

- For any w^t , the reported strategy up to time t can be written as $\sigma^t\left(w^t\right) = \left(\sigma_1\left(w^1\right), \sigma_2\left(w^2\right), ..., \sigma_t\left(w^t\right)\right)$.
- Note that $\sigma^T\left(w^T\right) = \sigma\left(w^T\right)$.

The welfare generated by the reporting strategy $\sigma(.)$ is equal to:

$$W(\sigma) = \int \left[\sum_{t=1}^{T} \beta^{t-1}\left[u\left(x_t\left(\sigma^t\left(w^t\right)\right)\right) - v\left(\frac{y_t\left(\sigma^t\left(w^t\right)\right)}{w^t}\right)\right]\right] d\mu_W$$

For the mechanism to be truthful, we must have:

$$W(\sigma^*) \geq W(\sigma) \text{ for all } \sigma$$

Planner's Problem

The planner's problem is to characterize the allocation $\left\{x_t\left(w^T\right), y_t\left(w^T\right)\right\}_{t=1}^{T}$ which **maximizes the expected lifetime utility of agents**:

$$\int \left[\sum_{t=1}^{T} \beta^{t-1}\left[u\left(x_t\left(w^t\right)\right) - v\left(\frac{y_t\left(w^t\right)}{w_t}\right)\right]\right] d\mu_W$$

subject to:

- The feasibility constraint (i.e. the resource constraint):

$$\int \left[\sum_{t=1}^{T} \frac{y_t\left(w^t\right) - x_t\left(w^t\right)}{(1+r)^{t-1}}\right] d\mu_W$$

- The incentive-compatibility constraints:

$$W(\sigma^*) \geq W(\sigma) \text{ for all } \sigma$$

Note: **Full commitment of the planner has been assumed**.

Optimal allocation

Suppose that $\left\{x_t^*\left(w^T\right), y_t^*\left(w^T\right)\right\}_{t=1}^T$ is the optimal allocation. Consider the following deviation at time t and after history $\overline{w}^t$ from this allocation:

$$u\left(x_t'\left(\overline{w}^t\right)\right) = u\left(x_t^*\left(\overline{w}^t\right)\right) + \varepsilon \tag{5}$$
$$u\left(x_{t+1}'\left(\overline{w}^t, w_{t+1}\right)\right) = u\left(x_{t+1}^*\left(\overline{w}^t, w_{t+1}\right)\right) - \beta^{-1}\varepsilon$$

- By construction, this new allocation yields the same utility as the optimal allocation for all w^T.
- It follows that the new allocation is also incentive-compatible.
- Note that, for ε sufficiently small:

$$u'\left(x_t^*\left(\overline{w}^t\right)\right) \simeq \frac{u\left(x_t^*\left(\overline{w}^t\right)\right) - u\left(x_t'\left(\overline{w}^t\right)\right)}{x_t^*\left(\overline{w}^t\right) - x_t'\left(\overline{w}^t\right)}$$
$$x_t^*\left(\overline{w}^t\right) - x_t'\left(\overline{w}^t\right) \simeq \frac{-\varepsilon}{u\left(x_t^*\left(\overline{w}^t\right)\right)}$$

from (5).

Optimal allocation

Similarly, we also have:

$$x^*_{t+1}\left(\overline{w}^t, w_{t+1}\right) - x'_{t+1}\left(\overline{w}^t, w_{t+1}\right) \simeq \frac{\beta^{-1}\varepsilon}{u'\left(x^*_{t+1}\left(\overline{w}^t, w_{t+1}\right)\right)}$$

The resources generated by the deviation are equal to:

$$x^*_t\left(\overline{w}^t\right) - x'_t\left(\overline{w}^t\right) + \frac{E\left[x^*_{t+1}\left(\overline{w}^t, w_{t+1}\right) - x'_{t+1}\left(\overline{w}^t, w_{t+1}\right)\middle| w^t = \overline{w}^t\right]}{1+r}$$

$$\simeq \frac{-\varepsilon}{u'\left(x^*_t\left(\overline{w}^t\right)\right)} + \frac{1}{1+r}E\left[\frac{\beta^{-1}\varepsilon}{u'\left(x^*_{t+1}\left(\overline{w}^t, w_{t+1}\right)\right)}\middle| w^t = \overline{w}^t\right]$$

Optimal allocation (cont'd)

$$\simeq \varepsilon \left(\frac{-1}{u'\left(x_t^*\left(\overline{w}^t\right)\right)} + \frac{1}{\beta\left(1+r\right)} \right)$$
$$+ \frac{1}{1+r} E \left[\frac{1}{u'\left(x_{t+1}^*\left(\overline{w}^t, w_{t+1}\right)\right)} \middle| w^t = \overline{w}^t \right]$$

Optimal allocation

If the initial allocation $\left\{x_t\left(w^T\right), y_t\left(w^T\right)\right\}_{t=1}^{T}$ is optimal, then the deviation must not generate any additional resources.
Thus, at the optimum, the term in brackets must be equal to zero.
This yields the **inverse Euler equation**:

$$\frac{1}{u'\left(x_t^*\left(\overline{w}^t\right)\right)} = \frac{1}{\beta\left(1+r\right)} E\left[\left.\frac{1}{u'\left(x_{t+1}^*\left(\overline{w}^t, w_{t+1}\right)\right)}\right| w^t = \overline{w}^t\right]$$

- Golosov, Kocherlakota and Tsyvinski (ReStud 2003) show that **this equation must hold for any stochastic process for wages**.
- This is the **cornerstone of the new dynamic public finance literature**.

Intuition for the Inverse Euler Equation

The inverse marginal utility of consumption has a straightforward interpretation:

- Let $X(u)$ denote the resource cost of providing utility u to a worker.
- Thus, by construction, $X(u(x_t^*)) = x_t^*$.
- Differentiating this expression with respect to c_t^* yields $X'(u(x_t^*))u'(x_t^*) = 1$.
- Hence, the inverse marginal utility of consumption $1/u'(x_t^*)$ is equal to the marginal resource cost of providing utility $X'(u(x_t^*))$.

It follows that the inverse Euler equation can be written as:

$$X'(u(x_t^*(\overline{w}^t))) = \frac{1}{\beta(1+r)} E\left[\frac{1}{u'\left(x_{t+1}^*(\overline{w}^t, w_{t+1})\right)} \middle| w^t = \overline{w}^t \right]$$

The planner allocates resources across time such as to minimize the resource cost of providing utility to workers.

Implications of the Inverse Euler Equation

In the absence of risk, or when idiosyncratic shocks are observable:

- Consumption at each point in time is independent of the shocks;
- The inverse Euler equation reduces to the standard Euler equation:

$u'(x_t^*) = \beta(1+r)u'(x_{t+1}^*)$
By Jensen, the function $x \longrightarrow 1/x$ being convex, for any positive random variable of positive variance Y:

$$\frac{1}{E\left(Y\right)} < E\left(\frac{1}{Y}\right)$$

Therefore, with hidden idiosyncratic shocks, by Jensen's inequality, we have:

$$E\left[\left.\frac{1}{u'\left(x^*_{t+1}\left(\overline{w}^t, w_{t+1}\right)\right)}\right| w^t = \overline{w}^t\right] > \frac{1}{E\left[1u'\left(x^*_{t+1}\left(\overline{w}^t, w_{t+1}\right)\right)\right| w^t = \overline{w}^t\right]}$$

which, by the inverse Euler equation, implies that:

$$u'\left(x^*_t\left(\overline{w}^t\right)\right) < \beta\left(1+r\right)E\left[u'\left(x^*_{t+1}\left(\overline{w}^t, w_{t+1}\right)\right)\right| w^t = \overline{w}^t\right]$$

as we had found in (4) with the two periods model.

Intuition

- If w_{t+1} is known for sure at date t, so is x_{t+1}, and the government should let the agent save and borrow at the before tax interest rate r (separable preferences).
- When w_{t+1} is random, compared with laissez faire present consumption is encouraged, savings is discouraged at the second best optimum. The existence of savings make it more costly for the government to provide incentives to work.

Implications of the Inverse Euler Equation

The failure of the standard Euler equation implies that:

- If agents could borrow and lend freely at the interest rate r, they would like to postpone consumption.
- Thus, to implement the optimal allocation, the planner must prevent agents from saving too much.

When the planner shifts resources into the future, to preserve incentives to work, he must increase the utility in the good states (with high skills at $t+1$) as much as in the bad states (with low skills at $t+1$).

- This requires a larger increase in consumption in the good state, where the marginal utility of consumption is low, than in the bad state, where it is high.
- Thus, the allocation of consumption across states will be less efficient at $t+1$ than at t.
- So, starting from the standard Euler equation, the planner wants to shift resources from $t+1$ to t.

The intertemporal wedge

The intertemporal wedge $t_{t+1}(.)$ induced by the optimal allocation is implicitly defined by:

$$u'\left(x_t^*\left(\overline{w}^t\right)\right) = \beta(1+r)\left(1-\tau_{t+1}\left(\overline{w}^t\right)\right) E\left[u'\left(x_{t+1}^*\left(\overline{w}^t, w_{t+1}\right)\right)\middle| w^t = \overline{w}^t\right]$$

Note that, by construction, $t_{t+1}(.)$ is w^t measurable.
The **intertemporal wedge** measures the discrepancy between:

- The marginal rate of substitution between consumption at t and at $t+1$;
- The marginal rate of transformation between production at t and at $t+1$.

We know that this wedge is positive, i.e. $\tau_{t+1}(.) > 0$ for any w^t such that w_{t+1} is random.

It is tempting to assert that $\tau_{t+1}\left(\overline{w}^t\right)$ is the optimal tax on wealth. We can show that this tax system does not implement the optimal allocation [I skip the proof]

- Setting the tax on savings equal to the wedge implements the optimal allocation **provided that agents keep supplying the optimal level of labor**.
- However, agents choose to follow a **double deviation** by **saving more in period in 1 and shirking in period 2** (even when their productivity is high).

From mechanism to taxes

The tax system must now operate in **two dimensions, labor supply and savings**. The above formula makes sure that savings is chosen as desired, once labor supply is optimal. Conversely the tax would implement the desired labor supply, given optimal savings. **They do not prevent joint deviations**.

Optimal taxes

How can the optimal allocation $\left\{x_t\left(w^T\right), y_t\left(w^T\right)\right\}_{t=1}^{T}$ be implemented in a decentralized economy using taxes which are a function of labor incomes and savings?

Kocherlakota (Ecta 2005) makes the wealth tax paid at $t+1$ conditional on the wage earned at $t+1$:

$$\tau_{t+1}^*\left(y^{t+1}\right) = 1 - \frac{u'\left(x^t\right)}{\beta\left(1+r\right)u'\left(x^{t+1}\right)}$$

where current consumption $x_t\left(w^t\right)$ depends on the history of skills w_t only through the history of labor incomes $y^t(w^t)$.

Now it can be shown that the choices of the optimizing agents follow from the usual Euler equation:

$$u'\left(x_t\left(w^t\right)\right)=\beta\left(1+r\right)E_t\left[\left(1-\tau^*_{t+1}\left(y^{t+1}\left(w^{t+1}\right)\right)\right)u'\left(x_{t+1}\left(w^{t+1}\right)\right)\right]$$

i.e. the inverse Euler equation holds!
Kocherlakota (Ecta 2005) shows under further assumptions that it is possible to design **an income tax** which together with the above wealth tax implements the optimum.

The optimal wealth tax is regressive!

How much revenue is raised from taxes on capital?
At any time t and for any history $\overline{w}^t$, it can be shown that the expected tax rate on savings is equal to zero! It brings zero tax receipts in the aggregate.
We now know that:

- The inverse Euler equation implies a positive wedge.
- The average rate of the tax on savings which implements the optimal allocation is equal to zero.

How can the tax on savings implement a positive intertemporal wedge even though it does not raise any revenue?

- The tax rate τ^*_{t+1} is high when income x^*_{t+1} is low and the marginal utility of consumption $u'\left(x^*_{t+1}\right)$ is high.
- The tax rate τ^*_{t+1} is low when income x^*_{t+1} is high and the marginal utility of consumption $u'\left(x^*_{t+1}\right)$ is low.

Thus, taxes on savings are high when savings are most needed for consumption smoothing which prevents the double deviation.

In other words, it discourages savings by making it a more risky investment: its rate is positive (resp. negative) when $1/u'(x_{t+1})$ is smaller (larger) than its expected value at t. The agent is taxed at a higher rate when her consumption is smaller.

Alternative Implementation

Although the optimal allocation of resources is unique, there are typically several ways to implement this allocation in a decentralized economy.

- We have just seen the implementation proposed by Kocherlakota (Ecta 2005).
- We have solved the implementation problem even though we have not fully characterized the optimal allocation of resources. Thus, **we cannot say anything about the optimal taxes on labor incomes.**

Dynamic public finance literature

The optimal allocation of resources could only be fully characterized when skills follow specific stochastic processes.

- Diamond Mirrlees (JPubE 1978) and Golosov Tsyvinski (JPE 2006):
    - Workers can be hit by a permanent disability shock which reduces their productivity to zero.
    - The optimal allocation can be implemented with an **asset test**, i.e. agents are only eligible to disability benefits if their assets are smaller than an age-specific threshold.
- Albanesi Sleet (ReStud 2006):
    - Skills are independently and identically distributed.
    - The optimal allocation can be implemented with a non-linear tax on both wealth and current labor income.

Dynamic public finance literature

- Farhi Werning (2013):
    - Skills follow a Markov process (such as AR(1) which is consistent with the empirical literature on wage shocks.
    - On average, labor wedges are increasing with age (from 0 to 37% over 40 years) while intertemporal wedges are decreasing (from 12 to 0% of the income from savings).
    - Farhi and Werning (2013) also showed that: Restricting taxes on labor income to be age-dependent and taxes on capital to be age-independent leads to a very small welfare loss (smaller than 0.09% of consumption). $\Rightarrow$ Allowing for history dependence adds a lot of complexity but little social value.
- Farhi and Werning (JPE 2012) managed to isolate the welfare gains from the inverse Euler equation in a general equilibrium setup:
    - **Switching from the standard Euler equation to the inverse Euler equation generates a welfare gain smaller than 0.2% of consumption.**
    - Thus, the inverse Euler equation does not provide a strong justification for taxing capital

Capital income inequality also due to other reasons

- **Shifting of labor income:** [skip]
  - The higher the shifting elasticity, the closer the tax rates on labor and capital income should be, Christiansen and Tuomala (ITAX 2008)
  - In practice, this seems to be an important consideration when designing tax systems, especially for top incomes.
- **Inheritances**: [skip]
  - Critical to understand why there are inheritances to decide on optimal inheritance tax policy.
  - 4 main models of bequests: (a) accidental, (b) bequests in the utility, (c) manipulative bequest motive, (d) dynastic.

Outline

Albanesi, S. and Sleet, C. (2006), Dynamic Optimal Taxation with Private Information, Review of Economic Studies, 73(1), 1–30.

Atkinson, A.B. and Stiglitz, J.E. (1976), The design of tax structure: Direct versus indirect taxation, Journal of Public Economics, 6, 55–75.

Banks, J. and Diamond, P. (2010), The base for direct taxationMirrlees Review. In James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba, eds., Dimension of tax design: The Mirrlees review, Oxford: Oxford University Press, 548-648.

Barro, R.J. (1974), Are government bonds net wealth?, Journal of Political Economy, 82, 1095-1117.

Bernheim, B. and Bagwell, K. (1988), Is everything neutral?, Journal of Political Economy, 96, 308-338.

**Boadway, R. (2012), From Optimal Tax Theory to Tax Policy, Retrospective and Prospective Views, CESifo Munich lectures in economics, MIT Press, Chapter 3.
Chamley, C. (1986), Optimal taxation of capital income in general equilibrium with infinite lives, Econometrica, 54(3), 607–622.
Chamley, C. (2001), Capital income taxation, wealth distribution and borrowing constraints, Journal of Public Economics, 79, 55–69.
Christiansen, V. and Tuomala, M. (2008), On taxing capital income with income shifting, International Tax and Public Finance, 15, 527–545.
Conesa, J.C., Kitao, S. and Krueger, D. (2009), Taxing capital? Not a bad idea after all!, American Economic Review, 99(1), 25–48.
Correia, I. (2010), Consumption taxes and redistribution, American Economic Review, 100, 1673–1694.

**Cremer, H. and Gahvari, F. (1995), Uncertainty, optimal taxation and the direct versus indirect tax controversy, Economic Journal, 105, 1165-1179.
Cremer, H., Pestieau, P. and Rochet, J.C. (2003), Capital income taxation when inherited wealth is not observable, Journal of Public Economics, 87, 2475–2490.
**Diamond, P. (2007), Comment on Golosov et al. NBER Macroeconomic Annual 2006, 365–379.
Diamond, P. and Mirrlees, J. (1978), A Model of Social Insurance with Variable Retirement, Journal of Public Economics, 10, 295–336.

Diamond, P. and Spinnewijn, J. (2011), Capital Income Taxes with Heterogeneous Discount Rates, American Economic Journal: Economic Policy, 3(4), 52–76.

Erosa, A. and Gervais, M. (2002), Optimal taxation in life-cycle economies, Journal of Economic Theory, 105, 338–369.

Farhi, E. and Werning, I. (2012), Capital Taxation: Quantitative Explorations of the Inverse Euler Equation, Journal of Political Economy, 120(3), 398-445.

Farhi, E. and Werning, I. (2013), Insurance and Taxation over the Life Cycle, Review of Economic Studies, 80(2), 596–635.

Farhi, E., Sleet, C., Werning, I. and Yeltekin, S. (2012), Non-linear capital taxation without commitment, Review of Economic Studies, 79(4), 1469–1493.

Fischer, S. (1980), Dynamic inconsistency, cooperation and the benevolent dissembling government, Journal of Economic Dynamics and Control, 2, 93–107.
Golosov, M., Troshkin, M., Tsyvinski, A. and Weinzierl, M. (2013), Preference heterogeneity and capital income taxation, Journal of Public Economics, 97,160–175.
Golosov, M., Kocherlakota, N. and Tsyvinski, A. (2003), Optimal indirect and capital taxation, Review of Economic Studies, 70, 569–597.
Golosov, M. and Tsyvinski, A. (2006), Designing Optimal Disability Insurance: A Case for Asset Testing', Journal of Political Economy, 114(2), 257–279.
Golosov, M. Tsyvinski, A. and Werning, I. (2007), New dynamic public finance: A user's guide, NBER, Macroeconomics Annual, 2006, 317–363.
Judd, K.L. (1985), Redistributive taxation in a simple perfect foresight model, Journal of Public Economics, 28, 59–83.

Kaplow, L. (1994), Taxation and risk taking: A general equilibrium perspective, National Tax Journal, 47(4), 789–798.

Kaplow, L. (2006), On the undesirability of commodity taxation even when income taxation is not optimal, Journal of Public Economics, 90, 1235–1250.

Kocherlakota, N. (2005), Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation, Econometrica, 73, 1587–1621.

Kocherlakota, N.R. (2010), The New Dynamic Public Finance, Princeton, NJ: Princeton University Press; chapter 2.

Konishi, H. (1995), A Pareto-improving commodity tax reform under a smooth non-linear income tax, Journal of Public Economics, 56, 413–446.

Krusell, P., Kuruscu, B. and Smith, A.A. (2010), Temptation and Taxation, Econometrica, 78(6), 2063–2084.

Laroque, G. (2005), Indirect taxation is superfluous under separability and taste homogeneity: A simple proof, Economics Letters, 87, 141–144.
*[Strongly recommended] Ljungqvist, L. and Sargent, T.J. (2012), Recursive macroeconomic theory, 3rd, Edition, Cambridge, MA: MIT Press; chapter 16, sections 2 to 8.
Ordover, J.A. and Phelps, E.S. (1979), The concept of optimal taxation in the overlapping-model of capital and wealth, Journal of Public Economics, 12, 1–26.
Persson, T. and Tabellini, G. (2002), Political Economics and Public Finance. In Alan J. Auerbach and Martin Feldestein, eds., Handbook of Public Economics, vol.3 Amsterdam: North Holland, 1549-1659.
Piketty, T. Saez, E. and S. Stantcheva (2014), Optimal Taxation of Top Incomes: A Tale of Three Elasticities, American Economic Journal: Economic Policy, 6 (1): 230-271

Saez, E. (2002), The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes, Journal of Public Economics,, 83, 217–230.
Straub, L. and Werning, I. (2014), Positive long run capital taxation: Chamley-Judd revisited, Working Paper, MIT.
Zucman, G. (2013), The Missing Wealth of Nations, Are Europe and the U.S. net Debtors or net Creditors?, Quarterly Journal of Economics, 128(3): 1321–1364.
Zucman, G. (JEP 2014), Taxing Across Borders: Tracking Personal Wealth and Corporate Profits, Journal of Economic Perspectives, 2014, 28(4): 121–148.