Les 6èmes

« TOULOUSE LECTURES IN ECONOMICS »

17-18-19 MARS 2008

Sixth Toulouse Lectures in Economics
March 17-19, 2008

The New Dynamic Public Finance

Narayana Kocherlakota

University of Minnesota, FRB-Minneapolis, and NBER

INSTITUT D'ECONOMIE INDUSTRIELLE

PARTENAIRE MAIRIE DE TOULOUSE www.toulouse.fr

PRINCETON

Lecture 2

Properties of Dynamic Optimal Taxation Systems

- A government has to make purchases over dates and states.

- It raises revenue via taxes on labor income and on wealth (and possibly other ways).

- What are the properties of the optimal taxes?

- In Lecture 1, I said that the new dynamic public finance says to follow two steps.

- Step 1: Find a socially optimal allocation conditional on skills and effort being private information.

- Step 2: Find a tax system that implements this socially optimal allocation.

- In Lecture 1: we did a lot of work on step 1.

- We derived important characteristics of a PO allocation, given private information about skills and effort.

- Recall: if (c, y) is PO, and $Var(u'(c_{t+1})|\theta^t) > 0$, then:

$$u'(c_t) < \beta R E_t u'(c_{t+1})$$

where c_t is individual consumption and y_t is individual output.

- Social planner needs to deter individual savings.

- In Lecture 2, we move away from the social planner's problem.

- We focus on the problem of how to design taxes to *implement* PO allocations.

Structure of Lecture 2

- I first write down a model economy with capital and aggregate shocks.

- I generalize the intertemporal Euler equation from Lecture 1.
  - as before, there is an intertemporal wedge

- It would seem natural to design optimal tax by setting it equal to this wedge.

- I use a 2 period example to show this approach does not work.

- Such a tax system is susceptible to a "double deviation".

- In a double deviation, agents save too much and then shirk in the following period.

- Any optimal period t wealth taxes must depend on period t labor income

- This finding is surprising: period t wealth is determined in period $(t-1)$.

- But optimal taxes are based on information in period t.

- This tax structure is what is needed to deter "double deviation": save and shirk

- Given a PO allocation, I show how to construct a tax system that implements it.

- In this system, taxes are linear in wealth.

- The optimal system has two main characteristics.

1. Average wealth tax rates (over people) are always zero.

2. Wealth tax rates are higher for people with surprisingly low labor income.

- Intuition?

- Zero average tax rates - basically uniform commodity taxation result.

- More important: "regressive" wealth taxes deter double deviation.

- By blocking "saving and shirking", wealth taxes allow for more social insurance.

- In the rest of the talk:

- I do a preliminary numerical exercise about wealth taxes.

- I use the NDPF approach to discuss:
  - optimal monetary policy
  - optimal debt policy

- Conclusions: private sector vs. public sector social insurance?

1. Environment

- Specification is similar to Lecture 1.

- I add capital and aggregate shocks.

- These additions allow us to address macroeconomic policy issues in a more realistic setting.

- There are T periods and a unit measure of agents.

- Agents have expected utility preferences, with utility function:

$$\sum_{t=1}^{T} \beta^{t-1}\{u(c_t) - v(l_t)\}, 0 < \beta < 1$$

where c_t is period t consumption and l_t is period t labor (measured in terms of effort).

- There are two types of shocks:
  - Private idiosyncratic shocks θ_t with finite support Θ.
  - Public aggregate shocks z_t with finite support Z.

Probability Structure

- Nature draws $(z_1, ..., z_T)$ from Z^T according to μ_Z.
- All agents learn z_t at the beginning of period t.

- Idiosyncratic shocks are determined as follows.

- Let z^{t+1} be the history of public shocks in period $(t+1)$.

- Let $\mu_\Theta(\theta'|\theta^t, z^{t+1})$ be a density over $\theta' \in \Theta$, given θ^t and z^{t+1}.

- Given public history z^{t+1}, and private history θ^t, Nature draws θ_{t+1} from $\mu_\Theta(.;\theta^t, z^{t+1})$

- These draws are iid across the agents with history θ^t.

- Given a public history z^t, the probability of having private history θ^t is given by:

$$\mu_\Theta(\theta^t|z^t) = \mu_\Theta(\theta_t;\theta^{t-1}, z^t)\mu_\Theta(\theta_{t-1};\theta^{t-2}, z^{t-1})...\mu_\Theta(\theta_1; z_1)$$

.

- We assume a LLN applies.

- Thus, $\mu_\Theta(\theta^t|z^t)$ is also the fraction of agents with private history θ^t, given public history z^t.

Economic Impact of θ's

- An agent with realization θ_t can produce y_t units of effective labor:

$$y_t = \theta_t l_t$$

- As before: y_t is public information.

- The skill θ_t and effort l_t are privately known to the agent.

Feasibility

- Let K_t denote per-capita capital; initial capital K_1^*.

- Let $(G_t)_{t=1}^T$, where $G_t : Z^t \longrightarrow R_+$, be required government purchases.

- An allocation is $(c_t, y_t, K_{t+1})_{t=1}^T$, where:

$$c_t : \Theta^t \times Z^t \longrightarrow R_+$$
$$y_t : \Theta^t \times Z^t \longrightarrow R_+$$
$$K_{t+1} : Z^t \longrightarrow R_+$$
$$K_1 \leq K_1^*$$

- Then, an allocation is feasible if for all t, z^t.

$$C_t(z^t) + K_{t+1}(z^t) + G_t(z^t)$$

$$\leq (1-\delta)K_t(z^{t-1}) + F_t(K_t(z^t), Y_t(z^t), z^t)$$

where:

$$C_t(z^t) \equiv \sum_{\theta^t \in \Theta^t} \mu_\Theta(\theta^t | z^t) c_t(\theta^t, z^t)$$
$$Y_t(z^t) \equiv \sum_{\theta^t \in \Theta^t} \mu_\Theta(\theta^t | z^t) y_t(\theta^t, z^t)$$

- Here, F_t is an aggregate production function.

Incentive-Compatibility and Optimality

- Information about θ is private.

- Hence, achievable allocations must also respect IC constraints.

- We can generate these as in Lecture 1 by using the Revelation Principle.

- A reporting strategy is $\sigma = (\sigma_t)_{t=1}^{T}$ where $\sigma_t : \Theta^t \times Z^t \to \Theta$.

- Let Σ be the set of all σ.

- Then, an allocation (c, y) is *incentive-compatible* if and only if for all σ in Σ

$$
\begin{aligned}
& E\{\sum_{t=1}^{T} \beta^{t-1}[u(c_t(\theta^t, z^t)) - v(\frac{y_t(\theta^t, z^t)}{\theta_t})]\} \\
\geq \; & E\{\sum_{t=1}^{T} \beta^{t-1}[u(c_t(\sigma^t(\theta^t, z^t), z^t)) - v(\frac{y_t(\sigma^t(\theta^t, z^t), z^t)}{\theta_t})]\}
\end{aligned}
$$

- As in Lecture 1, we are interested in period 1 Pareto Optima.

- Hence, the planner's objective may weight agents with different θ_1's differently.

Generality of the Shock Processes

- The model is agnostic about the time series behavior of z_t.

- The model is agnostic about the time series behavior of θ_t.

- The model allows for cyclical fluctuations in idiosyncratic risk.
  - $Var(\theta^{t+1}|\theta^t, z^{t+1})$ can depend on z_{t+1}.

- The model does not allow for private information about aggregates.
  - $\Pr(z^{t+1} = z^*|z^t, \theta^t)$ is independent of θ^t.

2. A Generalized "Reciprocal" Euler Equation

Theorem 1: Suppose (c, y, K) is a period 1 Pareto Optimum.

- Then, for all t, z^{t+1}, there exists $\lambda_{t+1}(z^{t+1})$ such that:

$$\lambda_{t+1}(z^{t+1}) = \frac{\beta E\{u'(c_{t+1}(\theta^{t+1}, z^{t+1}))^{-1}|\theta^t, z^{t+1}\}^{-1}}{u'(c_t(\theta^t, z^t))}$$

for all θ^t.

- And λ_{t+1} satisfies:

$$1 = E[\lambda_{t+1}(z^{t+1})(1 - \delta + F_{K,t+1}(z^{t+1}))|z^t]$$

for all t, z^t.

Content of Theorem 1

- $\lambda_{t+1}(z^{t+1})$ is the shadow discount factor of the planner.

- It equals the conditional *harmonic* mean of MU_{t+1} divided by MU_t.

- Note that: the denominator and numerator of RHS depend on θ^t.

- The theorem says that the ratio doesn't - and the ratio is the shadow discount factor of the planner.

- If Z is a singleton and $(1-\delta)+F_{K,t+1}(K_{t+1},Y_{t+1},z^{t+1})=R$, this result becomes:

$$1=\beta R\frac{E\{u'(c_{t+1}(\theta^{t+1}))^{-1}|\theta^t\}^{-1}}{u'(c_t(\theta^t))}$$

for all t,θ^t.

- This is the same as the "reciprocal" Euler equation in Lecture 1.

Intertemporal Wedge

- Suppose $Var(u'(c_{t+1})|\theta^t, z^{t+1}) > 0$ in some history θ^t, z^{t+1}.

- We expect this to be true in order to elicit effort from high types.

- Then, by applying Jensen's inequality, we can conclude that:

$$u'(c_t(\theta^t, z^t)) < \beta E[u'(c_{t+1}(\theta^{t+1}, z^{t+1})(1 - \delta + F_{K,t+1}(z^{t+1}))|z^t, \theta^t]$$

- As before, it's optimal to deter individual asset accumulation in some way.

3. Example: Taxes Don't Equal Wedges

- Until now: focus on Pareto optimal quantities.

- We now begin to think about taxes.

- We suppose people are able to buy and sell z^t-contingent assets, subject to taxes.

- How do we design an optimal tax structure?

- We shall see that using the natural ex-ante Pigouvian tax does not work.

Example Description

- Consider two-period example of the general environment.
- Set $\beta = 1$.
- Agents have period 1 endowment y; no production in period 1.
- Skill risk: agents are equally likely to have skills θ_H or θ_L in period 2.
- Let (c_i, y_i) be period 2 allocation of people with skills θ_i.
- Assume that: $F_2(K_2, Y_2) = RK_2 + Y_2$

The Ex-Ante Pigouvian Tax

- Suppose $(c_1^*, K_2^*, c_H^*, c_L^*, y_H^*, y_L^*)$ is optimal allocation, with $c_H^* > c_L^*$.

- Suppose it is an equilibrium given a tax rate τ on capital.

- We know from our earlier analysis:

$$u'(c_1^*) < R[u'(c_H^*) + u'(c_L^*)]/2$$

- Usual Pigouvian pub. fin. logic says to set the capital tax rate τ equal to this wedge.

- Here, this means setting τ so that:

$$(1 - \tau) = \frac{u'(c_1^*)}{R[u'(c_H^*) + u'(c_L^*)]/2}$$

Failure of the Ex-Ante Pigouvian Tax

- In the optimal allocation, the IC constraint binds:

$$u(c_H^*) - v(y_H^*) = u(c_L^*) - v(y_L^*/\theta_H)$$

- This means that, in the equilibrium, the agent is indifferent between two plans.

- Equilibrium plan
  - Save K_2^*
  - If $\theta = \theta_H$, set effort equal to y_H^*/θ_H
  - If $\theta = \theta_L$, set effort equal to y_L^*/θ_L

- Off-equilibrium plan of shirking when high-skilled
  - Save K_2^*
  - If $\theta = \theta_H$, set effort equal to y_L^*/θ_H.
  - If $\theta = \theta_L$, set effort equal to y_L^*/θ_L.

- But shirkers want to save more than K_2^*:

$$u'(c_1^*) = R(1-\tau)[u'(c_H^*)/2 + u'(c_L^*)/2]$$

$$\Rightarrow \; u'(c_1^*) < R(1-\tau)u'(c_L^*)$$

- So, there's a strictly better off-equilibrium plan: save AND shirk.
  - Save $K_2^* + \varepsilon$
  - Set effort equal to y_L^*/θ_i for $i = H, L$

- Under the ex-ante Pigouvian tax, the socially optimal allocation is not individually optimal.

- The individual is better off saving more and then shirking.

- The ex-ante Pigouvian tax does not implement optimal allocation.

4. An Optimal Wealth Tax System

- There exists an optimal tax system that is linear in wealth.
  - Tax rate on wealth differs across agents (based on current and past labor incomes).
  - Also: history-dependent taxes on labor income.

Optimal System: Ex-Post Pigouvian Taxes

- Go back to 2 period example.

- Set tax rate on capital equal to τ_H if $y_2 = y_H$ and τ_L if $y_2 = y_L$ where:

$$u'(c_1^*) = R(1 - \tau_i)u'(c_i^*), \;\; i = H, L$$

- Tax rate is set so that *ex-post* after-tax MRS equals the social MRT.

- Now, storing K_2^* is optimal even if the agent shirks.

- More generally, one can construct an optimal tax system with linear wealth taxes.

- In that optimal system, we set the wealth tax rate for an agent with skill history θ^{t+1} to be:

$$[1-\tau_{t+1}(\theta^{t+1},z^{t+1})]\frac{\beta u'(c^*_{t+1}(\theta^{t+1},z^{t+1}))}{u'(c^*_t(\theta^t,z^t))}=\lambda_{t+1}(z^{t+1})$$

where λ_{t+1} is as in Theorem 1.

- This tax system equates an agent's ex-post MRS with the planner's.
  - as in the two-period example

- It eliminates the desire to save more today, and then work less in the future.

Zero Average Wealth Taxes

- What's the expected tax rate, conditional on θ^t and z^{t+1}?

$$E(1-\tau_{t+1}(\theta^{t+1},z^{t+1})|\theta^t,z^{t+1})$$

$$= u'(c_t^*(\theta^t,z^t))\lambda_{t+1}(z^{t+1})\beta^{-1}E\{1/u'(c_{t+1}^*(\theta^{t+1})|\theta^t,z^{t+1}\}$$

- Theorem 1 says RHS $=1 \Rightarrow$ the expected tax rate is zero.

- For any agent at time t, the expected tax rate on his wealth at time $(t+1)$ is zero.

- Implies that the average tax rate across all agents is zero and that wealth tax collections are zero.

- Note: level of government purchases does not matter.

Optimal "Regressive" Wealth Taxes

- We set the optimal tax rate on wealth so that:

$$1 - \tau_{t+1}(\theta^{t+1}, z^{t+1}) = \frac{u'(c_t^*(\theta^t, z^t))}{\beta u'\left(c_{t+1}^*(\theta^{t+1}, z^{t+1})\right)} \lambda_{t+1}(z^{t+1})$$

- Given (θ^t, z^{t+1}), the tax rate is higher on the wealth of people with low $c_{t+1}^*(\theta^{t+1}, z^{t+1})$.

- Tax the wealth of unexpectedly low-skilled. Subsidize the wealth of unexpectedly high-skilled.

- No skill shocks - tax should be zero.

- Wealth tax seems regressive.

- But in total (with labor taxes), tax system transfers from high-skilled to low-skilled.

- The tax on wealth provides better incentives and so it allows the government to expand the transfers to low-skilled.

A Covariance Wedge

- How is a zero mean tax consistent with the optimal intertemporal wedge?

- Because of tax *risk*.

- The tax rate on wealth is negatively correlated with consumption.

- Optimal tax system deters investment by introducing risk to the agent.

Quantitative Properties of Optimal Wealth Taxes

- In Kocherlakota (2005), I solve for optimal wealth taxes in 2-period setting.

- Very preliminary exercise!

- I allow for persistent shocks to skills.

- There are three basic lessons of this exercise.

Lesson 1: The Need for Income Contingency

- The Ramsey literature argues that zero wealth taxes are optimal.

- But it abstracts from the optimal dependence of wealth tax rates on labor income.

- In my simple model, I compare two suboptimal systems.

- In the first, wealth taxes equal zero for all agents.

- In the second, wealth taxes depend optimally on labor income, except their average is required to be 2.5%.

- In the first system: too little social insurance.
  - because of saving and shirking problem
- In the second system: too little capital accumulation.
  - because average capital tax rates are too high
- Most tax analyses emphasize the deficiencies of the second system.
- I find that welfare may in fact be considerably lower in the first system.

Lesson 2: High Persistence Implies High Wealth Taxes/Subsidies

- Suppose the probability of transiting from θ_i to θ_j is low, $i \neq j$.

- Then, in the optimal wealth tax system ...

- The wealth tax after this transition is high in absolute value.

Lesson 3: If G is high, variance of optimal wealth taxes across agents is high.

- Intuition: if G is high, need lots of labor.
- Hence, economy need strong labor market incentives.
- It requires very high wealth taxes on the low-skilled.

5. Other Macroeconomic Policies

Monetary Policy: Transactions-Based Models

- Above, we studied a one good model.

- Suppose instead that agents receive utility from consuming two goods:

$$u(c_a, c_b) - v(y)$$

- Assume that c_a can be transformed one-for-one into c_b and vice-versa.

- What is the optimal tax rate on the different consumption goods?

- In any Pareto optimal allocation, for all (t, θ^t, z^t):

$$u_a(c_{at}(\theta^t, z^t), c_{bt}(\theta^t, z^t)) = u_b(c_{at}(\theta^t, z^t), c_{bt}(\theta^t, z^t))$$

- Suppose not, and $u_a > u_b$ in (θ^t, z^t).

- Then, there exists (c'_a, c'_b) such that:

$$u(c'_{at}(\theta^t, z^t), c'_{bt}(\theta^t, z^t)) \;=\; u(c_{at}(\theta^t, z^t), c_{bt}(\theta^t, z^t))$$

$$c'_{at}(\theta^t, z^t) + c'_{bt}(\theta^t, z^t) \;<\; c'_{at}(\theta^t, z^t) + c'_{bt}(\theta^t, z^t)$$

- Using (c'_a, c'_b) doesn't undo incentives - but gives more stuff to society.

- This shows that it's suboptimal to tax good a differently from good b.
- Now suppose c_a are cash goods and c_b are credit goods.
- Positive nominal interest rate = tax on cash goods.
- But such a tax is suboptimal - the nominal interest rate should equal zero.

- The above model is the cash-credit model of Lucas and Stokey (1987).

- However, the basic lesson is readily generalized to other monetary models like:
  - shopping-time models
  - money-in-the-utility function
  - Lagos-Wright matching model

- Key assumption: disutility of effort is separable from consumption.

Monetary Policy: Sticky Prices

- Modern monetary theory assumes changing prices is costly.
- Lessons from the Ramsey taxation literature: gov't can use profit or consumption taxes to undo the resulting distortions.
- Optimal monetary policy is still to set nominal interest equal to zero.
- Deviations from this policy are attributable to gov't not having profit or consumption taxes.

- The essence of Mirrlees taxation: government has enough instruments.

- Sticky prices don't change the basic message:

- It is still optimal to set the nominal interest rate equal to zero.

Debt Policy

- In this setting, optimal labor taxes and government debt are indeterminate.

- Essentially, the *timing* of when people pay taxes is irrelevant.

- Consider a two-period world, in which the production function is linear:

$$F(K, Y) = RK + Y$$

- Let (c^*, y^*, K^*) be a Pareto optimum in this world.

- Suppose too that there exists (c_1^+, c_2^+) such that:

$$c_1^+(y_1^*(\theta_1)) = c_1^*(\theta_1) \text{ for all } \theta_1$$

$$c_2^+(y_1^*(\theta_1), y_2^*(\theta_1, \theta_2)) = c_2^*(\theta_1, \theta_2) \text{ for all } (\theta_1, \theta_2)$$

- That is: consumption depends on θ only through y.

- Let $\tau^W(y_1, y_2)$ be the wealth tax rate in period 2 (as described above).
- Suppose we set labor income taxes (τ_1^L, τ_2^L) to be:

$$\tau_1^L(y_1) = y_1 - c_1^+(y_1) + K_1^* R - K_2^*(y_1)$$

$$\tau_2^L(y_1, y_2) = y_2 - c_2^+(y_1, y_2) + K_2^*(y_1) R(1 - \tau^W(y_1, y_2))$$

for all $(y_1, y_2) \in \{y_1^*(\theta_1), y_2^*(\theta_1, \theta_2)\}_{(\theta_1,\theta_2)\in\Theta^2}$.

- The labor taxes (τ_1^L, τ_2^L) and wealth taxes τ^W implement the Pareto optimal allocation.

- In this world, there is no government debt.

- Pick any function $\Delta(y_1)$.

- Now suppose we define:

$$\tau_1^{L*}(y_1) = \tau_1^L(y_1) - \Delta(y_1)$$

$$\tau_2^{L*}(y_1, y_2) = \tau_2^L(y_1, y_2) + \Delta(y_1)(1 - \tau^W(y_1, y_2))R$$

- The labor taxes $(\tau_1^{L*}, \tau_2^{L*})$ and wealth taxes τ^W also implement the Pareto optimal allocation.

- However, the path of government debt is now different.

- The government debt now equals the average of $\Delta(y_1)$ over all agents.

- This is a generalized debt irrelevance proposition.

- Suppose taxes are allowed to depend on past labor incomes.

- Then, given an optimal allocation, any path of debt is consistent with that allocation.

- In reality: work incentives are provided through a complicated package.
  - current period labor income taxes.
  - social security taxes and transfers.
  - disability insurance.
- There are many ways to re-structure these packages, without affecting incentives.
- All that's pinned down is the present value of payments, as a function of labor choices.
- Hence, optimal labor income tax schedules and debt policies are both indeterminate.

- One of the many optimal systems closely mimics US social security.
- Recall: the US social security system imposes a linear tax, on earnings below a ceiling.
- Then it pays a post-retirement transfer.
- The transfer depends on the entire lifetime labor income history.

- We can construct an optimal tax system with three components:
  - linear tax on labor income during working life
  - transfers during retirement, contingent on labor income history
  - linear, history-contingent, taxes on financial wealth as earlier

- Unlike current US system:
  - agents can borrow against transfers during retirement
  - wealth taxes depend on history of labor incomes

7. Conclusions

- We have discussed the structure of one optimal tax system.
- As I emphasized above, there are many optimal systems.
- Same "problem" occurs in Ramsey literature.
- Akin to portfolio theory: lots of ways to attain desired allocation.

- We assumed that all social insurance is done by the government.
- In real world, there's a mix of private and public sector provision.
- In fact, the US Constitution essentially guarantees a government role.
  - 14th amendment bans indentured servitude.
  - 16th amendment allows federal income taxes.
  - together: at least some social insurance must be done by gov't.
- What determines the split between public and private?
- Important - but very hard - question.