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

The Question

- 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?

Old Dynamic Public Finance

- Uses Ramsey approach.

- Taxes are restricted to be linear functions of current variables.

- Government attempts to minimize distortions generated by linearity.

- Weakness in Ramsey approach: ad hoc specification of tax instruments.

- Why restrict taxes to be linear?

- Why are taxes only functions of current variables?

New Dynamic Public Finance

- The literature is a dynamic generalization of Mirrlees.

- Mirrlees' approach based on two insights.

- Insight 1: major risk in life is *skill* risk.

- Some are born with the ability to generate income with relatively little effort. Others are not.

- In a dynamic setting: Some lose their ability to generate income (for example, due to mental illness or back injury). Others do not.

- Why don't societies perfectly insure their members against these major risks?

- Just tax everyone at 100% and share the proceeds evenly.

- If society can tell who is high-skilled and who is not: good system.

- Just command the high-skilled people to work hard.

- Insight 2: skills are often private information.

- A high-skilled person can choose to act like a low-skilled person.
  - A high IQ person can act like he/she has low IQ.
  - A person with a strong back can act like it's injured.
  - A person may fake depression or other mental illness.

- The even split plan is still possible.

- But it is no longer desirable.

- High-skilled will act like low-skilled.

- If skills are private information, tax system has to provide high-skilled people *incentives* to provide effort.

- A good tax system efficiently trades off incentives and insurance.

- The new dynamic public finance looks for an optimal tax system as follows.

- 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.

- The (old) Ramsey approach and the (new) Mirrlees approach are fundamentally different.

- *The Ramsey approach*
  - government cannot use lump-sum taxes.
  - key economic force: government tries to use linear taxes to mimic lump-sum taxation.

- *The Mirrlees approach*
  - government is free to use lump-sum taxes - chooses not to.
  - key economic force: trade-off between incentives and insurance.

Lecture 1: Foundations for Dynamic Social Contracting

Lecture 2: Properties of Dynamic Optimal Tax Systems

Lecture 3: Intergenerational Redistribution

Structure of Lecture 1

- Lecture 1 is about the first step of the New Dynamic Public Finance.

- Hence, there are no markets, prices, or taxes.

- It is only about the behavior of *quantities*.

- I consider two distinct economic *environments*.

- An environment is a specification of preferences, information, and technology.

- These objects are immutable and unalterable.

- They form the fundamental constraints on what society can achieve.

- The two environments have the same preferences and technology.
- Key element: individual skills evolve stochastically over time.
- The main results are valid for **any** data generation process for skills.

- First environment: skills are publicly observable.

- I characterize behavior of consumption of Pareto Optima.
  - intertemporal Euler equations
  - dependence on past
  - inequality and mobility

- Next: change the environment so that skills are private information.

- We see how to change the planner's problem.

- Then: I characterize Pareta Optimal consumption.
  - intertemporal Euler equations
  - dependence on past
  - inequality and mobility

Main Conclusions

- With public information:
  - PO displays little history dependence.
  - no change in inequality and zero mobility.

- With private information:
  - PO typically displays rich history dependence.
  - it is PO to regulate individual asset accumulation.
  - growing consumption inequality
  - short-run mobility, but no mobility in long run

2. Model Setup

- T periods; unit measure of agents.

- Preferences:

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

where c_t is consumption and l_t is effort.

Agent Heterogeneity

- Let Θ be a finite set.
- In period 0, Nature draws θ_T from Θ_T for each agent.
- The draws are iid across agents, according to pdf μ.
- LLN: Fraction of agents with history θ_t equals $\mu(\theta_t)$.
- Assume: $\mu(\theta_t) > 0$ for any finite t.
- Agent learns θ_t at beginning of period t.

Examples of μ

- $T = 2; \Theta = \{\theta^H, \theta^L\}$
- $\mu(\theta^1, \theta^2) = 0.25$ for all $(\theta^1, \theta^2) \in \Theta^2$
  - θ_t is i.i.d. over time.
- $\mu(\theta^i, \theta^j) = 0.36$ if $i = j$ and $= 0.14$ if $i \neq j$.
  - θ_t is positively autocorrelated.
- By playing with μ, we can get any time-series behavior for θ.

Economic Impact of Shocks

- Agent's period t output of consumption goods is given by $y_t = \theta_t l_t$.

- Here, l_t is effort which enters the utility function of agent.

- I will call θ_t the agent's period t skill.

- Government requires G_t units of per-capita consumption at each date to produce public goods.

- Society can freely borrow and lend at gross interest rate R.

- In Lecture 2, we'll endogenize R.

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

$$\begin{aligned} c_t &: \Theta^t \rightarrow R_+ \\ y_t &: \Theta^t \rightarrow R_+ \end{aligned}$$

- An allocation is **feasible** if:

$$\sum_{\theta^T \in \Theta^T} \mu(\theta^T) \sum_{t=1}^{T} R^{1-t}[c_t(\theta^t) + G_t - y_t(\theta^t)] \leq 0$$

- We use the LLN here - the fraction of people with history θ^T equals $\mu(\theta^T)$.

3. Public Information Period 1 Pareto Optima

- Assume that θ_t and l_t are publicly observable at the beginning of period t.

- Then, (c, y) is Pareto Optimal in period 1 iff, for some ϕ, it solves:

$$\max \sum_{\theta^T \in \Theta^T} \mu(\theta^T)\phi(\theta_1) \sum_{t=1}^{T} \beta^{t-1}[u(c_t(\theta^t)) - v(\frac{y_t(\theta^t)}{\theta_t})]$$

$$s.t. \quad \sum_{\theta^T \in \Theta^T} \mu(\theta^T) \sum_{t=1}^{T} R^{1-t}[c_t(\theta^t) + G_t - y_t(\theta^t)] \leq 0$$

$$c_t(\theta^t), y_t(\theta^t) \geq 0$$

- Here, $\phi(\theta_1)$ represents the planner's weights, based on period 1 skills.

Characterization through FONC

- For interior solutions, there exists λ (Lagrange Multiplier) such that:

$$\beta^{t-1}u'(c_t(\theta^t))\phi(\theta_1) = \lambda R^{1-t}$$

$$\beta^{t-1}v'(\frac{y_t(\theta^t)}{\theta_t})\frac{\phi(\theta_1)}{\theta_t} = \lambda R^{1-t}$$

- In deriving these, I've cancelled $\mu(\theta_t)$ (probability of θ_t) from both sides.

- In words: the marginal utility of consumption grows at a constant rate.

- MRS between consumption and labor is θ_t.

Dynamics of Public Information Pareto Optima

- Three key properties ...

- First, **Euler equation**:

$$u'(c_t(\theta^t)) = \beta R\{\frac{\beta^{-t}R^{-t}\lambda}{\phi(\theta_1)}\}$$

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

- At a PO allocation, an agent does not want to deviate by borrowing or lending.

- There is no need to regulate the agent's access to the outside asset market.

- Another Euler equation is satisfied:

$$\begin{aligned}\frac{1}{u'(c_t(\theta^t))} &= \beta^{t-1}R^{t-1}\lambda^{-1}\phi(\theta_1)\\ &= \beta^{-1}R^{-1}\sum_{\theta\in\Theta}\frac{\mu(\theta^t,\theta)}{\mu(\theta^t)}\beta^t R^t\lambda^{-1}\phi(\theta_1)\\ &= \beta^{-1}R^{-1}E\{\frac{1}{u'(c_{t+1}(\theta^{t+1}))}|\theta^t\}\end{aligned}$$

- This "reciprocal Euler equation" will be important later.

- Second, **limited history dependence.**

- In a PO allocation:

$$c_t(\theta^t) = u'^{-1}\left(\frac{\beta^{1-t} R^{-t} \lambda}{\phi(\theta_1)}\right)$$

- Regardless of dependence in μ, c_t depends only on θ_1 and t.

- There is no dependence on $(\theta_2, \theta_3, ..., \theta_t)$.

- Third, **inequality and mobility**

- Suppose $\beta R = 1$.

- Then, if (c, y) is PO, $c_t(\theta^t) = c^*(\theta_1)$ for all (t, θ^t).

- **Inequality stays the same** for all t.

- Also **zero consumption mobility**: consumption ranking always stays same.

3. Private Information Pareto Optima

- As discussed earlier: a person's skills are hard for others to observe.
- It's easy for the highly skilled to mimic the less skilled.
- Society must provide *incentives* to get the right amount of effort.

- We now suppose θ_t and l_t are privately known to agent.
  - But y_t is publicly observable.

- With private information, not all feasible allocations can be achieved.

- For example, consider the allocation:
  - $c_t(\theta^t) = c^*$ for all (t, θ^t)
  - $y_t(\theta^t) = \theta_t$ for all (t, θ^t).

- This allocation isn't achievable by society.

- High θ_t people will pretend to be low θ_t people..

- How do we figure out which allocations are actually achievable by society?

- This seems like a hard question.
- We want to find any allocation that could be an equilibrium outcome of any possible game.
- The answer lies in the Revelation Principle.
- Any equilibrium outcome of any game is a truth-telling equilibrium outcome of a *direct mechanism.*

- Mechanically, pick any allocation (c, y).

- Imagine that each period, agents report their θ realization to a planner.

- Planner gives the agent (c, y) based on his current and past reports.

- Then, an allocation (c, y) is an equilibrium outcome of some game iff it induces agents to report the truth.

- We call such allocations *incentive-compatible*.

Formal Definition of Incentive-Compatibility

- Define a reporting strategy to be $\sigma = (\sigma_t)_{t=1}^{T}$ where $\sigma_t : \Theta^t \to \Theta$.
- Let Σ be the set of all σ.
- A reporting strategy represents an agent's strategy in a direct mechanism.
- Then, an allocation (c, y) is *incentive-compatible* if and only if:

$$\sum_{\theta^T \in \Theta^T} \mu(\theta^T) \sum_{t=1}^{T} \beta^{t-1}[u(c_t(\theta^t)) - v(\frac{y_t(\theta^t)}{\theta_t})]$$

$$\geq \sum_{\theta^T \in \Theta^T} \mu(\theta^T) \sum_{t=1}^{T} \beta^{t-1}[u(c_t(\sigma^t(\theta^t))) - v(\frac{y_t(\sigma^t(\theta^t))}{\theta_t})] \text{ for all } \sigma \in \Sigma$$

Definition of Pareto Optimality

- An allocation (c, y) is Pareto Optimal in period 1 iff, for some ϕ, it solves:

$$\max \sum_{\theta^T \in \Theta^T} \mu(\theta^T)\phi(\theta_1) \sum_{t=1}^{T} \beta^{t-1}[u(c_t(\theta^t)) - v(\frac{y_t(\theta^t)}{\theta_t})]$$

$$s.t.\ (c, y) \text{ is feasible and IC}$$

$$c_t(\theta^t), y_t(\theta^t) \geq 0$$

- Note that there is no real connection between agents.
- Hence, the planner's problem is basically a standard principal-agent problem ....
- ... where ϕ determines agent's reservation utility.

- NOTE: It is very important that agents are atomless.

- With a unit measure of agents: cross-sectional distribution of θ's is common knowledge.

- With finite number: agents don't know this distribution.

- They use information in their allocations to make inferences.

- This feedback will make planner's problem much more difficult.

Property 1: Reciprocal Euler Equation

- Theorem: Suppose (c, y) is Pareto Optimal. Then for all t, θ^t:

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

- Notice: this reciprocal Euler equation was also satisfied with public information!

- Proof:

- Suppose (c, y) is PO, and pick $k \in R$.

- Then, construct another allocation c' so that:

$$u(c'_t(\theta^t)) + \beta u(c'_{t+1}(\theta^{t+1})) = u(c_t(\theta^t)) + \beta u(c_{t+1}(\theta^{t+1})) + k$$

for all θ^{t+1} and:

$$c'_s(\theta^s) = c_s(\theta^s) \text{ for } s \notin \{t, t+1\} \text{ and for all } \theta^s$$

- and so that ...:

$$\sum_{\theta^t \in \Theta^t} \mu(\theta^t) c'_t(\theta^t) + R^{-1} \sum_{\theta^{t+1} \in \Theta^{t+1}} \mu(\theta^{t+1}) c'_{t+1}(\theta^{t+1})$$
$$= \sum_{\theta^t \in \Theta^t} \mu(\theta^t) c_t(\theta^t) + R^{-1} \sum_{\theta^{t+1} \in \Theta^{t+1}} \mu(\theta^{t+1}) c_{t+1}(\theta^{t+1})$$

- Note: (c', y) is feasible.

- Also: ranking of reporting strategies is same - (c', y) is IC.

- Hence, if (c, y) is PO, it must be true that $(c_t, c_{t+1}, 0)$ solves:

$$\max_{c'_t, c'_{t+1}, k} k$$

$$s.t.$$

$$u(c'_t(\theta^t)) + \beta u(c'_{t+1}(\theta^{t+1})) = k + u(c_t(\theta^t)) + \beta u(c_{t+1}(\theta^{t+1})) \; \forall \theta^{t+1}$$

$$\sum_{\theta^t \in \Theta^t} \mu(\theta^t) c'_t(\theta^t) + R^{-1} \sum_{\theta^{t+1} \in \Theta^{t+1}} \mu(\theta^{t+1}) c'_{t+1}(\theta^{t+1})$$

$$= \sum_{\theta^t \in \Theta^t} \mu(\theta^t) c_t(\theta^t) + R^{-1} \sum_{\theta^{t+1} \in \Theta^{t+1}} \mu(\theta^{t+1}) c_{t+1}(\theta^{t+1})$$

- FONC:

$$u'(c_t(\theta^t))\sum_{\theta}\eta_{t+1}(\theta^t,\theta) = \mu(\theta^t)\gamma$$

$$\beta u'(c_{t+1}(\theta^{t+1}))\eta_{t+1}(\theta^{t+1}) = R^{-1}\mu(\theta^{t+1})\gamma$$

- $\eta_{t+1}(\theta^{t+1})$ is the multiplier on the first constraint (contingent on θ^{t+1})

- γ is the multiplier on the second constraint.

- Substituting for η_{t+1}/γ in the first FONC gives:

$$\begin{aligned}\frac{1}{u'(c_t(\theta^t))} &= \sum_\theta \frac{\eta_{t+1}(\theta^t,\theta)}{\mu(\theta^t)\gamma} \\ &= \sum_\theta \frac{\beta^{-1}R^{-1}\mu(\theta^t,\theta)}{\mu(\theta^t)u'(c_{t+1}(\theta^t,\theta))}\end{aligned}$$

- This proves the theorem. QED

Optimal Intertemporal Wedge

- The theorem says that:

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

- Typically, in a PO, $Var(u'(c_{t+1}(\theta^{t+1}))|\theta^t) > 0$.

- Why? Consumption depends on skills to elicit good effort from high types.

- Then, using the convexity of $1/x$, Jensen's inequality implies:

$$u'(c_t(\theta^t)) < \beta RE[u'(c_{t+1}(\theta^{t+1}))|\theta^t]$$

Intuition for the Wedge

- Why doesn't planner do a risk-free transfer from pd. t to pd. $(t+1)$?

- Such a transfer reduces agent's effort in period $(t+1)$...

- ... because of wealth effects in labor supply.

- The planner doesn't save as much as agent would like.

Property 2: History Dependence

- With private information, Pareto Optima exhibit rich history dependence.
- Suppose (c, y) is Pareto Optimal, and there exists θ^{t-1}, θ, and θ' such that:

$$c_t(\theta^{t-1}, \theta') > c_t(\theta^{t-1}, \theta)$$

$$\mu(\theta_{t+1}|\theta^{t-1}, \theta') = \mu(\theta_{t+1}|\theta^{t-1}, \theta) \text{ for all } \theta_{t+1}$$

- Then $c_{t+1}(\theta^{t-1}, \theta', \theta_{t+1}) \neq c_{t+1}(\theta^{t-1}, \theta, \theta_{t+1})$ for some θ_{t+1}.
- That is, c_{t+1} depends on θ_t, even if μ doesn't.

- Suppose not. Then:

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

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

$$= \frac{1}{u'(c_t(\theta^{t-1},\theta'))}$$

which is a contradiction.

- The reciprocal Euler equation generates history dependence in consumption.

Property 3: Inequality and Mobility

- Suppose $\beta R = 1$ and (c, y) is Pareto optimal.
- The reciprocal Euler equation implies that $1/u'(c_t)$ is a martingale:

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

- Basic probability says that:

$$Var(\frac{1}{u'(c_{t+1})}) = Var(E_t\frac{1}{u'(c_{t+1})}) + E(Var_t\frac{1}{u'(c_{t+1})})$$

- In a typical PO, c_{t+1} depends on θ_{t+1}.
- Hence, the second term on RHS is positive.

- It follows that, if $\beta R = 1$:

$$\begin{aligned} Var(\frac{1}{u'(c_{t+1})}) &> Var(E_t \frac{1}{u'(c_{t+1})}) \\ &= Var(\frac{1}{u'(c_t)}) \end{aligned}$$

- It is Pareto optimal for consumption inequality to grow over time.

- In the short run, there is consumption mobility, because of incentives.

- What about long run mobility?

- Suppose $\beta R = 1$, and (c, y) is Pareto optimal.

- Reciprocal Euler equation says:
  - $1/u'(c_t)$ is a positive martingale.
  - $E[1/u'(c_t)]$ is the same and finite for all $t < \infty$.

- The Martingale Convergence Thm implies that for almost all θ^∞ :

$$\lim_{t\to\infty} \left|\frac{1}{u'(c_t(\theta^t))} - \xi(\theta^\infty)\right| = 0$$

where $\xi : \Theta^\infty \longrightarrow R_+$.

- If $u'(\infty) = 0$, then there exists $c^* : \Theta^\infty \longrightarrow R_+$ such that:

$$\lim_{t\to\infty} |c^*(\theta^\infty) - c_t(\theta^t)| = 0$$

for almost all θ^∞.

- It follows that:

$$
\begin{aligned}
& \lim_{t\to\infty} \lim_{\tau\to\infty} |c_t(\theta^t) - c_{t+\tau}(\theta^{t+\tau})| \\
\leq \; & \lim_{\tau\to\infty} \lim_{t\to\infty} |c_t(\theta^t) - c^*(\theta^\infty)| + |c^*(\theta^\infty) - c_{t+\tau}(\theta^{t+\tau})| \\
= \; & 0
\end{aligned}
$$

for almost all θ^∞.

- In this sense, there is **zero long run consumption mobility**.

4. Conclusions

- Without priv. info., Pareto optimality has strong dynamic implications.
  - no distortions in asset accumulation
  - little history dependence in consumption
  - constant consumption inequality
  - zero mobility in consumption

- With private information, Pareto optimality implies:
  - intertemporal wedge
  - history dependence in allocations
  - growing inequality (if $\beta R = 1$)
  - zero mobility in the long run (if $\beta R = 1$ and $u'(\infty) = 0$)

- These results are valid for any data generation process of skills ...

- as long as c_{t+1} is not completely predictable at time t.

- We will be especially interested in the intertemporal wedge:

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

- Planner must deter individuals from saving as much as they would like.

- Intuition: unregulated asset accumulation leads to tendency to save and shirk ...

- The long-run impact of these regulations on asset accumulation can be severe.

- Suppose $\beta R = 1$, and agents can borrow and lend freely at interest rate R.

- Then, agent consumption obeys the Euler equation:

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

so that $u'(c_t)$ is a martingale.

- We can apply the martingale convergence theorem to show that:

$$\lim_{t\to\infty} c_t(\theta^t) = \infty$$

for almost all θ^∞.

- Intuition: agents accumulate infinite amounts of wealth for precautionary savings.

- In contrast, suppose $\beta R = 1$ and $u'(\infty) = 0$.

- Suppose too that (c, y) is Pareto optimal (with private info).

- Then $\lim_{t\to\infty} c_t(\theta^t) = c^*(\theta^\infty) < \infty$ almost everywhere.

- So: Unregulated savings leads to *infinite* long-run consumption ....

- ... while Pareto optimal consumption is *finite* in the long run.

- It is optimal to deter savings - a lot!
- How do we best accomplish this deterrence?
- That is the question that we take up in the next lecture.