Lecture Notes

New Dynamic Public Finance

Thorsten V. Koeppl

Queen's University

Fall Term 2014

1 A 2-period Moral Hazard Problem

1.1 Model

<u>Risk-neutral principal:</u>

—▷ pays a wage to an agent

—▷ discounts the future according to $\beta = \frac{1}{1+r}$

<u>Risk-averse agent:</u>

—▷ risk-averse agent: chooses action $a \in A$

—▷ preferences: $u(w) - c(a)$ with discounting according to β

—▷ outcomes: $\{x_1, \ldots, x_N\}$ with probability $\pi_s(a)$

—▷ strategy: $(s_0, s_1, \ldots s_N)$

<u>Contract z:</u>

—▷ state-contingent path of wages – or, equivalently, "pay-offs" $u(w)$ for the agent

—▷ finite N implies a contract is a list of $N + N^2$ wages

1.2 Main Results

Proposition 1.1. *The Pareto-optimal contract satisfies*

$$\frac{1}{u'(w_i)} = \sum_{j=1}^{N} \pi_j(s_i) \frac{1}{u'(w_{ij})} \tag{1.1}$$

for all $i = 1, \ldots, N$.

Proof. Consider any contract z and the optimal strategy s given z. Change the contract for only one state i in the first period to

$$\tilde{z}_i = z_i - y \tag{1.2}$$

$$\tilde{z}_{ij} = z_{ij} + \frac{y}{\beta} \text{ for all } j = 1, \dots, N. \tag{1.3}$$

Then, the original strategy is still optimal, since (i) for $j \neq i$ the contract doesn't change, (ii) the relative pay-offs for ij do not change and (iii) the NPV of the contract remains unchanged.

It must then be the case that the optimal contract minimizes the costs of the principal at $y = 0$. The costs for the principal is given by

$$u^{-1}(z_i + y) + \beta \sum_{j=1}^{N} \pi_j(s_i) u^{-1}\left(z_{ij} - \frac{y}{\beta}\right). \tag{1.4}$$

The FONC needs to be 0 at $y = 0$ which gives

$$\frac{1}{u'(w_i)} = \sum_{j=1}^{N} \pi_j(s_i) \frac{1}{u'(w_{ij})} \tag{1.5}$$

which completes the proof. □

Properties of the optimal contract

1. Memory: If $w_i \neq w_j$, then there exist k such that $w_{ik} \neq w_{jk}$.

2. "Martingale Property": If $\frac{1}{u'}$ is convex (concave/linear), $w_i \geq (\leq / =) \sum_{k=1}^{N} \pi_k(s_i) w_{ik}$.

"Savings-constrained" agent

The agent will not want to borrow if he bears income risk in the 2nd period. Instead, the agent would always like to save some of his wage in the first period for additional consumption in the second period.

Consider the problem

$$\max_b u(w_i - b) + \beta \sum_{j=1}^{N} \pi_j(s_i) u(w_{ij} + (1+r)b) \tag{1.6}$$

which yields a necessary condition equal to

$$-u'(w_i - b) + \sum_{j=1}^{N} \pi_j(s_i) u'(w_{ij} + (1+r)b) = 0. \tag{1.7}$$

At $b = 0$, this FONC must be positive, since we have from the optimal contract

$$u'(w_i) = \frac{1}{\sum_{j=1}^{N} \frac{\pi_j(s_i)}{u'(w_{ij})}} \leq \sum_{j=1}^{N} \pi_j(s_i) u'(w_{ij}) \tag{1.8}$$

where the last inequality follows from the weighted arithmetic mean being larger than the weighted harmonic mean (by Jensen's inequality).

Hence, with the optimal contract the agent would like to set $b > 0$ if given the opportunity to save.

2 Generalizing the Inverse Euler Equation

2.1 Model

- measure one of agents
- preferences

$$\sum_{t=1}^{T} \beta^{t-1} \left[u(c_t) - v(l_t) \right] \tag{2.1}$$

  where u strictly concave, v strictly convex and both are bounded
- idiosyncratic shocks: θ^T drawn from μ_Θ
- effective labour: $y_t(\theta^T) = \phi_t(\theta^T) l_t(\theta^T)$
- open economy: $\beta R = 1$

Assumptions:

1. People privately learn θ_t at the beginning of period t.

2. Output y_t and consumption c_t are publicly observed.

Hence, allocations in period t are only θ_t measurable.

Remark: Note that the agents can chose a particular (c, y), once they have observed their labour productivity ϕ. After reporting ϕ, the planner instructs them to deliver output y which is associated with utility $u(c(\phi)) - v(y(\phi)/\phi^*)$, where ϕ^* is the true realized idiosyncratic productivity shock.

2.2 Pareto Problem

Let ω be the utility level promised to a group of people. A Pareto optimal allocation (c^*, y^*) solves for some ω^*

$$\max_{c,y} \sum_{\theta^T} \sum_t \beta^{t-1} \mu(\theta^T) \left[u(c_t(\omega^*, \theta^T)) - v(y_t(\omega^*, \theta^T)/\phi_t(\theta^T)) \right] \tag{2.2}$$

subject to

$$\sum_{\theta^T} \sum_t \beta^{t-1} \mu(\theta^T) \left[u(c_t(\omega, \theta^T)) - v(y_t(\omega, \theta^T)/\phi_t(\theta^T) \right] \geq \omega \text{ for all } \omega \neq \omega^* \tag{2.3}$$

$$\sum_{\omega} \sum_{\theta^T} \sum_t R^{-t} \mu(\omega) \mu(\theta^T) \left[c_t(\omega, \theta^T) - y_t(\omega, \theta^T) \right] \leq 0 \tag{2.4}$$

$$V(\sigma_{TT}; c, y, \omega) \geq V(\sigma; c, y, \omega) \text{ for all } \sigma, \omega \tag{2.5}$$

The constraints are ex-ante promised utility, intertemporal feasibility and truthtelling, respectively.

Step 1 – Perturbation

Consider any incentive feasible allocation (c^*, y^*). Then, for some time t and some group with utility ω, change the allocation to (c', y^*) according to

$$u(c'_t(\omega^*, \theta^T)) = u(c^*_t(\omega^*, \theta^T)) + \Delta + \epsilon(\theta^t) \text{ for all } \theta^T \tag{2.6}$$

$$u(c'_{t+1}(\omega^*, \theta^T)) = u(c^*_{t+1}(\omega^*, \theta^T)) - \beta^{-1}\epsilon(\theta^t) \text{ for all } \theta^T \tag{2.7}$$

$$\sum_{\theta^T}[c'_t(\omega^*, \theta^T) - c^*_t(\omega^*, \theta^T)]\mu(\theta^T) + R^{-1}\sum_{\theta^T}[c'_{t+1}(\omega^*, \theta^T) - c^*_{t+1}(\omega^*, \theta^T)]\mu(\theta^T) = 0 \tag{2.8}$$

This perturbation is incentive feasible, since

- it leaves all other utilities ω untouched
- it scales utilities V by Δ for all reporting strategies σ
- it is resource feasible.

Step 2 – Pareto Problem Rewritten

The optimal allocation solves the problem

$$\max_{\Delta,\epsilon,c'_t,c'_{t+1}} \Delta \tag{2.9}$$

$$\text{subject to}$$

$$(2.6) - (2.8)$$

The solution must be $\Delta = 0$, $\epsilon = 0$, and $c' = c^*$.

Step 3 – FONC at $(0, 0, c^*_t, c^*_{t+1})$

Note that we are looking at specific dates t and $t+1$ for the perturbation which vary across *all paths with initial history* θ^t.

$$\sum_{\theta^T} \eta(\theta^T) = 1 \tag{2.10}$$

$$-\sum_{\theta^T \geq \theta^t} \eta_t(\theta^t) + \beta^{-1}\sum_{\theta_{t+1}}\sum_{\theta^T \geq (\theta_{t+1},\theta^t)} \eta_{t+1}(\theta^t) = 0 \tag{2.11}$$

$$u'(c^*_t(\theta^T)) \sum_{\theta^T \geq \theta^t} \eta_t(\theta^T) = \lambda \sum_{\theta^T \geq \theta^t} \mu(\Theta^T) \tag{2.12}$$

$$u'(c^*_{t+1}(\Theta^T)) \sum_{\theta^T \geq (\theta_{t+1},\theta^t)} \eta_{t+1}(\theta^T) = \lambda R^{-1} \sum_{\theta^T \geq (\theta_{t+1},\theta^t)} \mu(\Theta^T) \tag{2.13}$$

Rewriting, we obtain the result

$$\frac{1}{u'(c_t^*(\omega^*, \Theta^T))} = E\left[\frac{1}{u'(c_{t+1}^*(\omega^*, \Theta^T))}|\theta^t\right] \tag{2.14}$$

where we have used the fact that $\beta R = 1$.

—▷ The inverse of the marginal utility follows thus a martingale. Any change on the inverse of marginal utility today has the same expected change on the inverse of marginal utility in the future. Hence, all shocks have permanent effects.

—▷ Why does it work? The key here is that both consumption and the marginal utility of consumption are publicly observable for the planner. This allows us to use the perturbation method as in Rogerson (1985) to characterize Pareto-optimal allocations.

—▷ Again, we have that there is a wedge in the standard Euler equation,

$$u'(c_t^*(\omega^*, \Theta^T)) < E\left[u'(c_{t+1}^*(\omega^*, \Theta^T))|\theta^t\right] \tag{2.15}$$

which implies that people are savings-constrained.

—▷ The idea here is that it is better to reduce smoothing of consumption (a second-order loss) for better insurance today (a first-order benefit).

3 Dynamic Mirrlees Taxation

3.1 General Idea

Ramsey Taxation:

- planner needs to use linear taxes
- minimize distortions (deadweight loss) from linear taxes
- cannot choose lump-sum taxes

Mirrless Taxation:

- planner can choose any tax system he wants
- but faces frictions (information, enforcement, etc.)
- optimal tax system achieves a *constrained* Pareto optimal allocation
- need to balance insurance vs. incentives
- can choose lump-sum taxes, but does not want to

3.2 Model

- measure one of agents
- preferences
$$\sum_{t=1}^{T} \beta^{t-1} \left[u(c_t) - v(l_t)\right] \tag{3.1}$$
where u strictly concave, v strictly convex and both are bounded
- aggregate shock: z^T drawn from μ_Z
- idiosyncratic shocks: θ^T drawn from μ_Θ
- aggregate shock z_t and θ_t learned at the beginning of period t
- effective labour: $y_t(\theta^T, z^T) = \phi_t(\theta^T, z^T) l_t(\theta^T, z^T)$
- effective labour is publicly observed; labor input and skills are private information
- aggregate production function CRS

Assumption: Again all shocks are drawn at the start of time. Hence, all variables in period t are functions of the shocks drawn, but are measurable only with respect to the history of shocks revealed up to period t.

3.3 The Inverse Euler Equation Once More

Feasible allocation:

$$\sum_{\theta^T} c_t(\theta^T, z^T)\mu(\theta^T) + K_{t+1}(z^T) + G(z^T) \leq F(K_t, Y_t, z^T) + (1-\delta)K_t(z^T) \tag{3.2}$$

where $Y_t(z^T) = \sum_{\theta^T} y_t(\theta^T, z^T)\mu_{\Theta^T}$ and $G(z^T)$ is government expenditure.

Incentive Compatability:

—▷ strategy: $\sigma : \theta^T \times Z^T \to \theta^T \times Z^T$

—▷ pay-off: $V(\sigma; c, y) = \sum_{t=1}^{T} \beta^{t-1} \sum_{z^T} \sum_{\theta^T} \left[u(c_t(\sigma)) - v(l_t(\sigma))\right] \mu(\theta^t)\mu(z^t)$

—▷ truthtelling strategy σ^*

An allocation is incentive compatible, if

$$V(\sigma^*; c, y) \geq V(\sigma; c, y) \tag{3.3}$$

for all σ.

A Pareto-optimal allocation maximizes ex-ante expected utility subject to being resource feasible and incentive compatible.

We again use the fact that there cannot be any way to redistribute consumption between today and tomorrow's states to save costs, while leaving the expected utility of any agent the same at any point in time for any shock (θ^T, z^T) – which implies incentive compatibility.

We solve a perturbed problem given by

$$\min_{c_t, c_{t+1}, K_{t+1}, \xi} \sum_{\theta^T} c_t(\theta^T)\mu(\theta^T) + K_{t+1} \tag{3.4}$$

subject to

$$u(c_t(\theta^T)) = u(c_t^*(\theta^T, z^t)) + \beta \sum_{z_{t+1}} \xi(\theta^T, z_{t+1})\mu(z_{t+1}|z^t) \text{ for all } \theta^T \tag{3.5}$$

$$u(c_{t+1}(\theta^T, z_{t+1})) = u(c_{t+1}^*(\theta^T, z_{t+1})) - \xi(\theta^T, z_{t+1}) \text{ for all } \theta^T, z_{t+1} \tag{3.6}$$

$$\sum_{\theta^T} c_{t+1}(\theta^T)\mu(\theta^T) - F_{t+1}(K_{t+1}, Y_{t+1}(z^t), z^t) - (1-\delta)K_{t+1} = -K_{t+2}(z_{t+1}, z^t) - G_{t+1} = \text{ for all } z_{t+1} \tag{3.7}$$

The first-order necessary conditions are given by

$$\mu(\theta^t) - \eta_t(\theta^t)u'(c_t) = 0 \tag{3.8}$$

$$-u'(c_{t+1})\eta_{t+1}(\theta_{t+1}) + \gamma(z_{t+1}|z^t)\mu(\theta^{t+1}) = 0 \text{ for all } z_{t+1} \tag{3.9}$$

$$1 - \sum_{z_{t+1}} \gamma(z_{t+1}|z_t)\left[1 - \delta + MPK(z_{t+1}|z^t)\right] \tag{3.10}$$

$$\beta\eta_t(\theta^t)\mu(z_{t+1}|z^t) - \sum_{\theta_{t+1}} \eta_{t+1}(\theta^{t+1}) = 0 \text{ for all } z_{t+1} \tag{3.11}$$

where – slightly abusing notation – η's and μ's are understood to be the sum of all probabilities across future paths given a history θ^t. Define $\lambda_{t+1} = \frac{\gamma(z^{t+1}|z^t)}{\mu(z_{t+1}|z^t)}$ which yields the following result.

Proposition 3.1. *Suppose* (c^*, y^*, K^*) *is an optimal allocation. Then, there exists a* z_{t+1}*-measurable function* $\lambda^*_{t+1} : Z^T \to R_+$ *such that*

$$\lambda^*_{t+1} = \beta\frac{1}{E\left[\frac{u'(c^*_t)}{u'(c^*_{t+1})}|\theta^t, z^{t+1}\right]} \tag{3.12}$$

$$E\left[\lambda^*_{t+1}(1 - \delta + MPK(z_{t+1}|z^t)|z^t)\right] = 1 \tag{3.13}$$

Again, we get a wedge in the intertemporal Euler equations. To see this, use first Jensen's inequality to obtain

$$\lambda^*(z_{t+1}) < \beta E\left[\frac{u'(c^*_{t+1})}{u'(c^*_t)}|\theta^t, z^{t+1}\right] \tag{3.14}$$

for all z_{t+1} succeeding z^t. Plugging into the second equation and using the law of iterated expectations, we obtain

$$\beta E\left[u'(c^*_{t+1})(1 - \delta + F_{Kt+1})|\theta^t, z^t\right] > u'(c^*_t). \tag{3.15}$$

Below for implementing the optimal allocation, we need to make sure that the intertemporal Euler equation holds with equality in the decentralized economy.

3.4 Interpreting λ^*_{t+1}

We call λ^*_{t+1} the *social discount factor*.

—▷ The Lagrange multiplier λ^*_{t+1} is the shadow value of a unit of more resources tomorrow. It expresses the discounted value of an additional amount of resources next period in event z_{t+1} taking into account the probability of the event. The shadow value of today's resources has been normalized to 1.

—▷ The proposition states that the social discount factor is equal to the harmonic mean of the MRS conditional on θ^t and is independent of individual histories θ^t. That is all agent's harmonic mean of the MRS has to be equal to $\lambda^*_{t+1}(z_{t+1})$ after history z^t.

—▷ The social discount factor then determines how much capital should optimally be accumulated.

—▷ The intuition for this result is that an extra unit of consumption needs to be split in such a fashion as to keep the utility level (!) fixed across different histories θ^t. This is very different from raising everyone's consumption by a fixed amount in such a fashion as to equate marginal utilities.

3.5 Decentralization through a Tax System

We restrict ourselves to

- non-linear labour taxes $\psi : I\!R^T_+ \times Z^T \rightarrow I\!R^T$
- linear capital taxes $\tau : I\!R^T_+ \times Z^T \rightarrow I\!R^T$

Hence, the agent pays taxes *ex post* on new and old capital according to $\tau_t(y(\theta^t, z^t), z^t)(1 - \delta + r_t(z^t))k_t(\theta^t, z^t)$ where I have slightly abused notation with respect to states.

<u>Intuition:</u>

- It seems like a constant linear tax on capital can equate the Euler equation. But this is NOT incentive compatible.
- Why? Saving more and lying tomorrow about one's ability beats saving the right amount and telling the truth tomorrow.

- The idea is to increase the consumption risk by levying wealth (or here, capital) taxes that vary with one's ability – and, hence, income – thereby deterring savings.

How do capital taxes look like?

Set wealth (capital) taxes ex-post(!) such as to equate agent's after-tax MRS with the social discount factor, or

$$\tau_{t+1}(y^T, z^T) = 1 - \lambda^*_{t+1}(z^T)\frac{u'(c^*_t(y^T, z^T))}{\beta u'(c^*_{t+1}(y^T, z^T))}. \tag{3.16}$$

Note that taxes depend on *observable* output and not directly on the announcement of skills. I assume here that there is a 1-1 mapping between the two.

—▷ Taxes are history-dependent (θ^t matters) and are *state-contingent* as they have to depend on next periods labour income through consumption. Hence, there is uncertainty for the household about the future capital tax rate.

—▷ Wealth taxes are high when future consumption is low and vice versa. This deters a deviation which includes saving more, work too little when skilled and claim to be unskilled tomorrow.

Results:

1) At these taxes, the intertemporal Euler equation of the agent is satisfied for the optimal allocation.

$$\begin{aligned}&\beta E[(1-\tau_{t+1})u'(c^*_{t+1})(1-\delta+r_{t+1}|\theta^t, z^t] - u'(c^*_t)\\ &= \left[E[\lambda^*_{t+1}(1-\delta+r_{t+1})|\theta^t, z^t] - 1\right] u'(c^*_t) = 0\end{aligned}$$

2) Labour taxes are lump-sum and thus can be chosen to satisfy the budget constraints at the optimal allocation. Hence, a 2nd Welfare Theorem holds.

3) Conditional on (θ^t, z^{t+1})), tomorrow's *expected* individual wealth tax is zero.

$$E[(1-\tau^*_{t+1}|\theta^t, z^{t+1}] = \lambda^*_{t+1}\beta^{-1}u'(c_t)E\left[\frac{1}{u'(c_{t+1})}|\theta^t, z^{t+1}\right]$$

4) Aggregate wealth taxes are zero for any history z^{t+1}.

$$\sum_{\theta^T} \tau^*_{t+1} k^*_{t+1} (1 - \delta + MPK^*_{t+1}) \mu(\theta^T) =$$
$$= (1 - \delta + MPK^*_{t+1}) E[\tau^*_{t+1} k^*_{t+1} | z^{t+1}] = (1 - \delta + MPK^*_{t+1}) E[E[\tau^*_{t+1} | \theta^t, z^{t+1}] k^*_{t+1} | z^{t+1}] = 0$$

Hence, capital taxes do not raise revenue and are purely redistributive.

5) Current wealth taxes are a decreasing function of people's consumption/skills (see above).

There are some issues: (i) government debt is irrelevant; (ii) agents cannot engage in side trades; (iii) the tax system is indeterminate.

4 Literature

Rogerson, Econometrica (1985)

Kocherlakota, Econometrica (2005)

Kocherlakota, The New Dynamic Public Finance (2010)