

New Dynamic Public Finance: Optimal Taxation with Private Information

Noah Williams

University of Wisconsin - Madison

Economics 714

The Question

- A government has to make purchases over dates and states.
- It raises revenue via taxes on labor income and on wealth.
- 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

- General Approach: Be explicit about the informational and/or enforcement frictions that limit government's extraction power.
- Find tax systems that generate optimal allocations, given those frictions.

- What kinds of frictions?
- So far, 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 (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.

Structure of Talk

I lay out a dynamic model economy in which:

- Skills are private information
- Effort is private information
- Skills can follow any stochastic process

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 a socially optimal allocation.

I discuss three results from using the new two-step approach.

1. intertemporal characterization of socially optimal allocations.
2. a necessary condition for optimal wealth taxes.
3. a description of wealth taxes in a particular optimal tax system.

- The results are all about wealth taxes.
- The optimal tax system also has labor income taxation.
- We know little about the specific structure of these labor income taxes.
- But they play an important role in both redistribution and revenue collection.

1. A General Environment

- T periods.

- Preferences:

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

where c_t is period t consumption and l_t is period t effort.

- Private idiosyncratic shock θ .
- (Note: no *aggregate* shocks in talk - but all results can be extended to the case of *public* aggregate shocks.)

Probability Structure of Shocks

- For each agent, θ^T is drawn from Borel set Θ^T according to prob. measure μ_Θ .
- $(\theta_1, \dots, \theta_T)$ is iid across agents.
- Information revelation: at the beginning of period t , each agent learns his θ_t .

Economic Impact of Shocks

- An agent's *effective* labor y_t is given by:

$$y_t = \theta_t l_t$$

- l_t is effort and θ_t is skill
- (l_t, θ_t) is private information; y_t is observable.

Feasibility

- Let K_t be per-capita capital; initial capital K_1^* .
- An allocation is $(c_t, y_t, K_{t+1})_{t=1}^{\infty}$, where:

$$c_t : \Theta^t \rightarrow R_+$$

$$y_t : \Theta^t \rightarrow R_+$$

$$K_{t+1} \in R_+$$

$$K_1 \leq K_1^*$$

- An allocation is feasible if for all t :

$$\begin{aligned} & \int_{\theta^t \in \Theta^t} c_t(\theta^t) d\mu_{\Theta} + K_{t+1} + G_t \\ & \leq (1 - \delta)K_t + F_t(K_t, \int_{\theta^t \in \Theta^t} y_t(\theta^t) d\mu_{\Theta}) \end{aligned}$$

- F_t is a CRS production function
- G_t is non-negative gov't purchases.

Incentive Compatibility

- Informational constraints imply that achievable allocations must be also be incentive-compatible.
- We find the incentive-compatible allocations using the revelation principle.
- Fix an allocation (c, y) . Each period, agents report their θ realization to a planner.
- Planner gives the agent (c, y) based on his current and past reports.
- The incentive-compatible allocations (c, y) are the ones that induce agents to report the truth.

- An allocation that is both incentive-compatible and feasible is said to be *incentive-feasible*.
- An *optimal* allocation is the incentive-feasible allocation that provides maximal ex-ante utility to agents.
- But results generalize to other notions of social optimality.

General Features of Environment

- Results are valid for any data generation process for skills.
- Desirable because there is an ongoing empirical debate about dynamics of individual wages.

Special Features of Environment

- Additive separability between consumption and leisure.
- Unskilled = more disutility from working.

2. Three Results and Their Implications

Statement of Result 1

In an optimal allocation, for all θ^t :

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

- Why does the RHS have conditional *harmonic* mean?
- Suppose planner gets Δ extra units of consumption in period $(t+1)$.
- Result 1 says that the planner's marginal benefit from these resources is:

$$\Delta[E(u'(c_{t+1}^*(\theta^{t+1}))^{-1}|\theta^t)]^{-1} \quad \text{NOT} \quad \Delta[E(u'(c_{t+1}^*(\theta^{t+1})|\theta^t)]$$

- The second expression would be right if the planner split extra consumption evenly across agents.
- But: incentive constraint implies the planner cannot split Δ evenly across agents.

.

- Instead, planner divides Δ into $\Delta(\theta_{t+1})$ to each person θ_{t+1} , so that each person's gain B is the same.

$$\Delta_{t+1}(\theta_{t+1})u'(c_{t+1}(\theta_{t+1})) = B$$

for all θ_{t+1} .

- B is also the planner's gain. But how big is B ?

$$\begin{aligned}\Delta_{t+1}(\theta^{t+1}) &= B/u'(c_{t+1}(\theta^{t+1})) \\ \Delta &= E(B/u'(c_{t+1})) \\ B &= \Delta[E(1/u'(c_{t+1}))]^{-1}\end{aligned}$$

- The planner's marginal benefit is given by $[E\{1/u'(c_{t+1})\}]^{-1}$ not $E\{u'(c_{t+1})\}$.

Optimal Intertemporal Wedge

- Result 1 says:

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

- Private information about θ_{t+1} implies that it is efficient for

$$\text{Var}(u'(c_{t+1}^*(\theta^{t+1}))|\theta^t) > 0.$$

- Jensen's inequality implies:

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

- It is optimal for the individual shadow interest rate to be lower than the social rate of return.

Intuition for the Wedge

- Risk-free savings from period t to period $(t + 1)$ reduce incentive to provide effort in period $(t + 1)$.
- optimal to deter such savings.

Statement of Result 2

- Suppose agents trade in a competitive equilibrium subject to wealth and labor income taxes.
- Suppose too that the equilibrium allocation is socially optimal.
- Then: the period t marginal tax rate on wealth W_t at the *beginning* of period t depends on effective labor y_t *during* period t .
- Optimal tax on an *ex-ante* decision (W_t) depends on *ex-post* outcome y_t .

Why Does Result 2 Work?

- 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: $F_2(K_2, Y_2) = RK_2 + Y_2$

- Suppose $(c_1^*, K_2^*, c_H^*, c_L^*, y_H^*, y_L^*)$ is optimal allocation. Then the IC constraint binds:

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

- Suppose the optimal allocation $(c_1^*, K_2^*, c_H^*, c_L^*, y_H^*, y_L^*)$ is an equilibrium given a tax schedule $\tau(K_2, y_2)$.

- In the equilibrium, the binding IC implies that 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 .

- Suppose Result 2 is wrong: the marginal tax rate τ_K is independent of y_2 .
- But then shirkers want to save more than K_2^* :

$$u'(c_1^*) = R(1 - \tau_K(K_2^*)) [u'(c_H^*)/2 + u'(c_L^*)/2]$$

$$\Rightarrow u'(c_1^*) < R(1 - \tau_K(K_2^*)) u'(c_L^*)$$

- So, there's a strictly better off-equilibrium plan: shirk AND save.
 - Save $K_2^* + \varepsilon$
 - Set effort equal to y_L^*/θ_i for $i = H, L$
- Result 2 follows: If optimal allocation is an equilibrium, then τ_K must depend on y_2 .

Statement of Result 3

- 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.

How Does Result 3 Work?

- 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:

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

- 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, the optimal wealth tax rate for an agent with skill history θ^{t+1} is:

$$1 - \tau_{t+1}(\theta^{t+1}) = \frac{u'(c_t^*(\theta^t))}{\beta(1 - \delta + F_{K,t+1}^*)u'(c_{t+1}^*(\theta^{t+1}))}$$

- What's the expected tax rate, conditional on θ^t ?

$$E(1 - \tau_{t+1}(\theta^{t+1})|\theta^t) = \frac{u'(c_t^*(\theta^t))}{\beta(1 - \delta + F_{K,t+1}^*)} E\{1/u'(c_{t+1}^*(\theta^{t+1})|\theta^t)\}$$

- Result 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.

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

$$1 - \tau_{t+1}(\theta^{t+1}) = \frac{u'(c_t^*(\theta^t))}{\beta(1 - \delta + F_{K,t+1}^*)u'(c_{t+1}^*(\theta^{t+1}))}$$

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

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.

Three Lessons from Dynamic Mirrlees Approach

I. optimal intertemporal wedge

II. the optimal marginal tax rate on savings has to depend on current labor income.

III. there exists an optimal tax system with linear tax on wealth; average wealth tax rate is zero.

Key Economic Point

Good social insurance requires higher taxes on savings of recipients.

The tax deters people from saving a lot, and then going on social insurance.

Allows for larger social insurance.