Dynamic Mirrlees Taxation: New Dynamic Public Finance

Mirrlees Approach

- So far representative agents or observable heterogeneity
- We have considered distortionary taxes but limited to be linear
- In Ramsey approach not deep theoretical justification of the use of distortionary taxes, while with Mirrlees approach distortionary taxes are optimal
- We now introduce **idiosyncratic uncertainty** with **private information**: Skills shocks or preference shocks
- SP maximizes ex-ante welfare and can use any potential nonlinear taxes
- SP must provide incentives for truth-telling $\Rightarrow$ Adverse selection as in Contract Theory

Mirrlees Approach

- As in contract theory, the Gvt acts as a Principal and tax-payers as Agents
- The presence of asymmetric information generates a **trade-off between insurance and incentives**
- Since tax instruments are not restricted, first best allocation would be attainable without heterogeneity
  - ⇒ Source of distortion comes from need to redistribute and insure in the presence of heterogeneity
  - ⇒ Taxes wedges will depend on skills distribution, risk aversion, labor supply elasticity

The Way Forward

1) Characterization of Pareto optimal consumption when skills are unobservable under assumption that Revelation Principle is satisfied
    - ◇ Inverse Euler Condition (Diamond and Mirrlees, 1978; Rogerson, 1985; Golosov, Kocherlakota, Tsvinsky, 2003)
2) Decentralization of optimal allocation through taxes when Gvt has full commitment

Framework

- There is a continuum of (ex-ante) identical agents living $T = 2$ periods (Young and Old):

$$u(c_1) + v(n_1) + \beta[u(c_2) + v(n_2)]$$

  with c consumption and n work effort, **additive separable**

- Output produced by each agent is:

$$\underbrace{y}_{\text{observable}} = \underbrace{\theta \cdot n}_{\text{unobservable}}$$

  with $\theta \in \Theta$ individual skills, which evolves stochastically according to π_Θ

- Each agent privately learns of θ after the realization
- Distribution of skills iid across agents
- Since agents are ex-ante identical but exposed to risk, only insurance no redistribution issues

Framework

- For simplicity consider a discrete number of realization of θ
- At $t = 1$, each agent draws $\theta_1(i)$, with $i = 1, 2, ..., N$ and denote by $\pi_1(i)$ the ex-ante probability of realization $\theta_1(i)$
- At $t = 2$, each agent draws $\theta_2(i, j)$, with $i, j = 1, 2, ..., N$ and denote by $\pi_2(j|i)$ the probability of realization $\theta_2(i, j)$ conditional on $\theta_1(i)$
- There can be persistence (Markov shocks) or not (intertemporal iid shocks)

Framework

- Production technology is linear in efficiency units of labor and saving technology is also linear
- There can be aggregate shocks: In $t = 2$ the rate of return of capital R_2 and public good expenditure G_2 are subject to shocks $s \in S$ which realize with prob $\mu(s)$
- The **resource constraints** are:

$$\sum_i \pi_1(i)[c_1(i) - y_1(i)] \leq R_1 K_1 - K_2 - G_1$$

$$\sum_{i,j} \pi_1(i)\pi_2(i,j)[c_2(i,j) - y_2(i,j)] \leq R_2(s) K_2 - G_2(s)$$

Step 1: Characterization Pareto Optimal Allocation

- To find the PO allocation, maximize welfare subject to information constraint
- Think of a fictitious SP who collects individual reports on productivity and allocate consumption and provide income conditional on the ability agents claim to have
- Each agent reporting strategies are functions $i_r(i)$ and $j_r(j,s)$
- Since skills are private information, the PO allocation must be such that no workers has incentive to mimic other types

Step 1: Characterization Pareto Optimal Allocation

The **Incentive Compatibility** constraint is given by:

$$
\begin{aligned}
& u(c_1(i)) + v\Big(\frac{y_1(i)}{\theta_1(i)}\Big) + \beta \sum_{j,s} \Big[u(c_2(i,j,s)) + v\Big(\frac{y_2(i,j,s)}{\theta_2(i,j)}\Big)\Big] \pi_2(i,j)\mu(s) \\
& \geq u(c_1(i_r(i))) + v\Big(\frac{y_1(i_r(i))}{\theta_1(i)}\Big) \\
& + \beta \sum_{j,s} \Big[u(c_2(i_r(i), j_r(j,s), s)) + v\Big(\frac{y_2(i_r(i), j_r(j,s), s)}{\theta_2(i,j)}\Big)\Big] \pi_2(i,j)\mu(s)
\end{aligned}
$$

Step 1: Characterization Pareto Optimal Allocation

- The constrained efficient planning problem maximizes expected discounted utility:

  $$\sum_i \Big(u(c_1(i)) + v\Big(\frac{y_1(i)}{\theta_1(i)}\Big) \\ + \beta \sum_{j,s} \Big[u(c_2(i,j,s)) + v\Big(\frac{y_2(i,j,s)}{\theta_2(i,j)}\Big)\Big] \pi_2(i,j)\mu(s)\Big)\pi_1(i)$$

  subject to the resource constraints and the incentive compatibility constraints

- Let (c^*, y^*, K^*) denote the solution to this problem

Step 1: Characterization Pareto Optimal Allocation

- If SP can observe private skills (i.e., max welfare only s.t. resource constraint), the optimality conditions are:

$$\Delta_{y_1}(i) \equiv 1 + \frac{v'\left(\frac{y_1^*(i)}{\theta_1^*(i)}\right)}{\theta_1(i)u'(c_1^*(i))}$$

$$\Delta_{y_2}(i,j,s) \equiv 1 + \frac{v'\left(\frac{y_2^*(i,j)}{\theta_2^*(i,j)}\right)}{\theta_2(i,j,s)u'(c_2^*(i,j,s))}$$

$$\Delta_k(i) \equiv 1 - \frac{u'\left(c_1^*(i)\right)}{\beta \sum_{j,s} \mu(s)\pi_2(j|i)R(s)u'(c_2^*(i,j,s))}$$

- Δ_y are intra-temporal wedges and Δ_k is the intertemporal wedge (Euler equation), all equal to zero in the first-best case

Benchmark: No Uncertainty in $t = 2$ (Atkinson-Stiglitz, 1976)

- Suppose that skills are constant over time, that is, $\theta_1(i) = \theta_2(i, j) = \theta(i)$, no aggregate risk, $R_1 = R_2(s) = R = 1 + r$, no $G_1 = G_2(s) = 0$

  $\Rightarrow$ All uncertainty is solved in $t = 1$

- **AS theorem** shows that if utility is weakly separable (intertemporally and intratemporally) and no uncertainty in $t = 2$, no need to tax capital

Benchmark: No Uncertainty in $t = 2$ (Atkinson-Stiglitz, 1976)

- Consider an even simpler case with only two types $\theta(2) > \theta(1)$ and labor supply only in $t = 1$
- Simplify notation: $x_t(\theta(i)) \equiv x_t^i$
- SP problem is:

$$\max_{c,y} \pi^1 \Big(u(c_1^1) + v\Big(\frac{y^1}{\theta^1}\Big) + \beta u(c_2^1) \Big) + \pi^2 \Big(u(c_1^2) + v\Big(\frac{y^2}{\theta^2}\Big) + \beta u(c_2^2) \Big)$$

s.t. RC and IC^2 (type 2 has incentives to mimic type 1)

$$\pi^1 \Big(c_1^1 + \frac{c_2^1}{1+r} - y^1 \Big) + \pi^2 \Big(c_1^2 + \frac{c_2^2}{1+r} - y^2 \Big) \leq RK_1 \quad (\lambda)$$

$$u(c_1^2) + v\Big(\frac{y^2}{\theta^2}\Big) + \beta u(c_2^2) \geq u(c_1^1) + v\Big(\frac{y^1}{\theta^2}\Big) + \beta u(c_2^1) \quad (\gamma \pi^2)$$

Benchmark: No Uncertainty in $t = 2$ (Atkinson-Stiglitz, 1976)

FOCs are:

$$\begin{aligned}
c_1^1 :& u'(c_1^1)(\pi^1 - \gamma\pi^2) = \lambda\pi^1 \\
c_2^1 :& u'(c_2^1)\beta(\pi^1 - \gamma\pi^2) = \lambda\pi^1 \frac{1}{1+r} \\
c_1^2 :& u'(c_1^2)(1+\gamma) = \lambda \\
c_2^2 :& u'(c_2^2)\beta(1+\gamma) = \lambda\frac{1}{1+r} \\
y^1 :& \frac{1}{\theta^1} v'\left(\frac{y^1}{\theta^1}\right) + \lambda - \gamma\frac{\pi^2}{\pi^1}\frac{1}{\theta^2} v'\left(\frac{y^1}{\theta^2}\right) = 0 \\
y^2 :& \frac{1}{\theta^2} v'\left(\frac{y^2}{\theta^2}\right) + \lambda + \gamma\frac{1}{\theta^2} v'\left(\frac{y^2}{\theta^2}\right) = 0
\end{aligned}$$

Benchmark: No Uncertainty in $t = 2$ (Atkinson-Stiglitz, 1976)

- Rearranging we get:

$$\begin{aligned} u'(c_1^1) = \beta(1+r)u'(c_2^1) &\Rightarrow \Delta_k(1) = 0 \\ u'(c_1^2) = \beta(1+r)u'(c_2^2) &\Rightarrow \Delta_k(2) = 0 \\ -v'\left(\frac{y^2}{\theta^2}\right) = \theta^2 u'(c_1^2) &\Rightarrow \Delta_y(2) = 0 \\ -v'\left(\frac{y^1}{\theta^1}\right) < \theta^1 u'(c_1^1) &\Rightarrow \Delta_y(1) > 0 \end{aligned}$$

- Same incentives as in a static problem: zero capital taxes, distortionary taxes only on low skill to provide incentives to high skill to not mimic low skill (regressive taxes, standard in contract theory)

Benchmark: No Uncertainty in $t = 2$ (Atkinson-Stiglitz, 1976)

- Capital taxes is equivalent to differential taxation of present and future consumption
- Without uncertainty on future productivity, both types choose a bundle of consumption (c_1^i, c_2^i) that smooth intertemporal utility (Euler condition holds)
- If the two types have the same disposable income, they choose the same consumption profile, which minimizes the resource cost of providing the equilibrium utility (dual approach)

Benchmark: No Uncertainty in $t = 2$ (Atkinson-Stiglitz, 1976)

- Furthermore, we have an intra-temporal wedge between types:

$$\frac{u'(c_t^1)}{u'(c_t^2)} = \frac{(1+\gamma)\pi^1}{\pi^1 - \gamma\pi^2}$$

- If γ is larger, then the spread in marginal utility is larger and more distortion on the low type is needed to provide incentives

- The allocation can be implemented with a combination of lump-sum transfers and non-linear (income) taxes, such that, the decentralized individual max program is:

$$\max u(c_1) + v(n_1) + \beta u(c_2)$$
$$\text{s.t. } c_1 + \frac{c_2}{1+r} \leq \theta n_1 (1-\tau) + T$$

Extensions and Robustness (Atkinson-Stiglitz, 1976)

AS theorem robust to

- Labor decisions also in the second period $\Rightarrow$ Differential commodity taxes on first and second period consumption
- Productivity vary over time (but not uncertainty in the second period) and agents commit to reporting in $t = 1 \Rightarrow$ Age-dependent income taxes
- Different discount factors, for example, $\beta^2 > \beta^1$ implies that intertemporal wedges is distorted for low type and need to tax capital to low skilled/low β (Saez, 2002; Diamond and Spinnewijn, 2011)

Inverse Euler Equation

- If there is uncertainty also in $t = 2$, then the intratemporal wedge will not be zero and therefore need to tax capital (Golosov, Tsyvinski and Werning, 2007)
- For any realization of i in $t = 1$ we must minimize the resource cost of delivering the utility from consumption
- Consider small perturbations Δ such that intertemporal utility remains unchanged:

$$\begin{aligned} u(\tilde{c}_2(i, j; \Delta)) &\equiv u(c_2(i, j)) + \Delta \\ u(\tilde{c}_1(i; \Delta)) &\equiv u(c_1(i)) - \beta\Delta \end{aligned}$$

Inverse Euler Equation

- If there is uncertainty also in $t = 2$, then the intratemporal wedge will not be zero and therefore need to tax capital (Golosov, Tsyvinski and Werning, 2007)
- For any realization of i in $t = 1$ we must minimize the resource cost of delivering the utility from consumption
- Consider small perturbations Δ such that intertemporal utility remains unchanged:

$$\begin{aligned} u(\tilde{c}_2(i,j;\Delta)) &\equiv u(c_2(i,j)) + \Delta \\ u(\tilde{c}_1(i;\Delta)) &\equiv u(c_1(i)) - \beta\Delta \end{aligned}$$

- Since this perturbation does not changes welfare and IC, but only RC, then $\{c\}$ is optimal (it minimizes resource cost), iff $\Delta = 0$ reaches the minimum

Inverse Euler Equation

- The resource cost to be minized is

$$\min_{\Delta} u^{-1}(u(c_1(i)) - \beta\Delta) + \frac{1}{R_2}\sum_j \pi(j|i)u^{-1}(u(c_2(i,j)) + \Delta)$$

- The FOC evaluated at $\Delta = 0$ yields the **Inverse Euler Equation**

$$\frac{\partial u^{-1}(\cdot)}{\partial c_1(i)}(-\beta) + \frac{1}{R_2}\sum_j \pi(j|i)\frac{\partial u^{-1}(\cdot)}{\partial c_2(i,j)} = 0$$

$$\frac{1}{u'(c_1(i))} = \frac{1}{\beta R_2}\sum_j \pi(j|i)\frac{1}{u'(c_2(i,j))}$$

$$\frac{1}{u'(c_1(i))} = \frac{1}{\beta R_2}\mathbb{E}_i\Big[\frac{1}{u'(c_2(i,j))}\Big]$$

Inverse Euler Equation

- Without uncertainty AS theorem applies

$$\frac{1}{u'(c_1(i))} = \frac{1}{\beta R_2}\frac{1}{u'(c_2(i))}$$
$$\Rightarrow$$
$$u'(c_1(i)) = \beta R_2 u'(c_2(i))$$

- With uncertainty, by the Jensen inequality, i.e., $\mathbb{E}[X] > \left(\mathbb{E}\left[\frac{1}{X}\right]\right)^{-1}$ (i.e., arithmetic mean>harmonic mean), the standard Euler equation must necessarily be distorted:

$$\begin{aligned}\frac{1}{u'(c_1(i))} &= \frac{1}{\beta R_2}\mathbb{E}_i\left[\frac{1}{u'(c_2(i,j))}\right] \\ &> \frac{1}{\beta R_2}\left[\frac{1}{\mathbb{E}_i[u'(c_2(i,j))]}\right] \Rightarrow \tau_k(i) \geq 0\end{aligned}$$

Inverse Euler Equation

- There is a double deviation incentive in place
- Agents who have incentive to mimic other types (high skill incentive to mimic low skill), have also incentive to deviate on saving
- Those deviating agents would like to increase saving for precautionary motive, which further increases their incentives to mimic low skill in the second period
- Taxing capital reduces the amount of consumption available in $t = 2$ from saving and thus reduces incentives to mimic low skill

Inverse Euler Equation

With aggregate risk, we obtain similar result (Kocherlakota, 2005)

$$\frac{1}{u'(c_1(i))} = \frac{1}{\beta}\frac{1}{\sum_s \left[R_2(s)\Big(\sum_j \pi(j|i)\frac{1}{u'(c_2(i,j,s))}\Big)^{-1}\right]\mu(s)}$$

$$\Rightarrow$$

$$u'(c_1(i)) < \beta \sum_{j,s} \mu(s)\pi_2(j|i)R(s)u'(c_2(i,j,s))$$

Step 2: Implementation Through Taxes

- It is tempting to assert that the intertemporal wedge $\Delta_k(i)$ is the optimal tax on wealth
- Setting the tax on savings equal to the wedge implements the optimal allocation provided that agents keep supplying the optimal level of labor
- Conversely the intratemporal wedge $\Delta_y(i)$ would implement the desired labor supply, given optimal savings
- 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)

Step 2: Implementation Through Taxes

- In the implementation stage, we consider the following decentralized maximization problem:

$$\max\left\{u(c_1(i)) + v\left(\frac{y_1(i)}{\theta_1(i)}\right) + \beta\sum_{j,s}\left(u(c_2(i,j,s)) + v\left(\frac{y_2(i,j,s)}{\theta_2(i,j)}\right)\right)\pi_2(j|i)\mu(s)\right\}$$

- subject to individual budgets:

$$c_1(i) \leq y_1(i) - k(i) - T_1(y_1(i))$$
$$c_2(i,j,s) \leq y_2(i,j,s) + R_2(s)k(i) - T_2(y_1(i), y_2(i,j,s), k(i), s)$$

Step 2: Implementation Through Taxes

- The socially optimal allocation $(c_1^*(i), c_2^*(i,j,s), y_1^*(i), y_2^*(i,j,s))$ is implementable if this allocation solve the individual maximization problem given the taxes $T_1(y_1(i))$ and $T_2(y_1(i), y_2(i,j,s), k(i), s)$
- Generally, a given socially optimal allocation can be implemented by various tax systems
- Simplest method to implement the optimal allocation is through **direct mechanism**, that is, imposing high punishments if chosen allocation is different from the optimal one, but unrealistic!
- Large literature on optimal tax systems which resemble real one

Step 2: Implementation Through Taxes

- Albanesi and Sleet (2006) show that the optimal taxes depends on labor supply and capital stock in that period $T_t(y_t, k_t)$, but only if skill shocks are iid and no aggregate risk
- Kocherlakota (2005) different implementation with distict taxes on labor and capital robust to various stochastic processes:
  - Labor taxes depend non-linearly on history of $y_t(i)$
  - Linear and history dependent capital tax which satisfies:

$$\tau_k(i,j) = 1 - \frac{u'(c_1^*(i))}{\beta R_2 u'(c_2^*(i,j))} \quad \text{with} \quad \sum_j \pi(j|i)\tau_k(i,j) = 0$$

- Since taxes are random, net return on savings are negatively correlated with income (risk premium), which reduces benefits from saving

Numerical Example: Setting

- Suppose $\theta_1 = 1$ and $\theta_2(1) = \theta > 1$ and $\theta_2(2) = 1/\theta$ with equal prob.
- No aggregate shocks, no public spending
- First best problem

$$\max u(c_1) + v(y_1) + \beta\Big(\frac{u(c_h) + v\left(\frac{y_h}{\theta}\right)}{2} + \frac{u(c_l) + v\left(\frac{y_l}{\frac{1}{\theta}}\right)}{2}\Big)$$

subject to:

$$c_1 + \frac{1}{R}\Big(\frac{c_h + c_l}{2}\Big) \leq y_1 + \frac{1}{R}\Big(\frac{y_h + y_l}{2}\Big) \quad (\lambda)$$

- Consider $u(c) = \log(c)$ and $v(n) = -\frac{n^2}{2}$ and $\beta = R = 1$

Numerical Example: First Best

- The FOCs from the first-best problem are:

$$
\begin{aligned}
c_1 : \frac{1}{c_1} - \lambda = 0 \\
c_h : \frac{1}{c_h} - \lambda = 0, c_l : \frac{1}{c_l} - \lambda = 0 \\
y_1 : -y_1 + \lambda = 0 \\
y_h : -\frac{y_h}{\theta^2} + \lambda = 0, y_l : -y_l\theta^2 + \lambda = 0 \\
c_1 + \left(\frac{c_h + c_l}{2}\right) = y_1 + \left(\frac{y_h + y_l}{2}\right)
\end{aligned}
$$

- Hence, $c_1 = c_2 = c_3$ (full insurance) and $n_h > n_l$: Is this allocation incentive compatible?
- No, since high skill has incentives to mimic low skill

Numerical Example: Incentive Compatibility

- If there is misreporting, agents get (c_l, n_l) and IC is:

$$\log(c_1) - \frac{y_1^2}{2} + \Big(\frac{\log(c_h) - \left(\frac{y_h}{\theta}\right)^2 \frac{1}{2}}{2} + \frac{\log(c_l) - \left(\theta y_l\right)^2 \frac{1}{2}}{2}\Big)$$

$$\geq \log(c_1) - \frac{y_1^2}{2} + \Big(\frac{\log(c_l) - \left(\frac{y_l}{\theta}\right)^2 \frac{1}{2}}{2} + \frac{\log(c_l) - \left(\theta y_l\right)^2 \frac{1}{2}}{2}\Big)$$

which implies

$$\log(c_h) - \log(c_l) \geq \frac{y_h^2 - y_l^2}{2\theta^2}$$

Numerical Example: Constrained Maximization

The constrained efficient problem is:

$$\max \log(c_1) - \frac{y_1^2}{2} + \Big(\frac{\log(c_h) - \frac{\left(\frac{y_h}{\theta}\right)^2}{2}}{2} + \frac{\log(c_l) - \left(\frac{(\theta y_l)^2}{2}\right)}{2}\Big)$$

subject to:

$$c_1 + \left(\frac{c_h + c_l}{2}\right) \leq y_1 + \left(\frac{y_h + y_l}{2}\right) \quad (\lambda)$$

$$\log(c_h) - \log(c_l) \geq \frac{y_h^2 - y_l^2}{2\theta^2} \quad (\gamma)$$

Numerical Example: FOCs

The FOCs of the constrained efficient problem are:

$$\begin{aligned}
c_1 &: \frac{1}{c_1} - \lambda = 0 \\
c_h &: \frac{1}{2c_h} - \lambda\frac{1}{2} + \gamma\frac{1}{c_h} = 0 \\
c_l &: \frac{1}{2c_l} - \lambda\frac{1}{2} - \gamma\frac{1}{c_l} = 0 \\
y_1 &: -y_1 + \lambda = 0 \\
y_h &: -\frac{y_h}{2\theta^2} + \lambda\frac{1}{2} - \gamma\frac{y_h}{\theta^2} = 0 \\
y_l &: -\frac{\theta^2 y_l}{2} + \lambda\frac{1}{2} + \gamma\frac{y_l}{\theta^2} = 0
\end{aligned}$$

Numerical Example: FOCs

- The consumption levels are:

$$c_1 = \frac{1}{\lambda}$$
$$c_h = \frac{1+2\gamma}{\lambda}$$
$$c_l = \frac{1-2\gamma}{\lambda}$$

- There is a positive intertemporal wedge if $\gamma > 0$:

$$\frac{c_h^*}{c_1^*} = 1+2\gamma \quad \frac{c_l^*}{c_1^*} = 1-2\gamma$$
$$\Rightarrow \Delta_k = 1 - \frac{1}{\frac{1}{2}\frac{c_1^*}{c_h^*} + \frac{1}{2}\frac{c_1^*}{c_l^*}} > 0$$

Numerical Example: FOCs

- The labor supplies are:

$$y_1 = \lambda$$

$$y_h = \frac{\lambda\theta^2}{1+2\gamma} < y_h^{FB}$$

$$y_l = \frac{\lambda\theta^2}{\theta^4 - 2\gamma} > y_l^{FB}$$

- The intratemporal wedges are:

$$\Delta_{y_1} = 1 - \frac{y_1^*}{\frac{1}{c_1^*}} = 0$$

$$\Delta_{y_2(h)} = 1 - \frac{\frac{y_h^*}{\theta}}{\frac{\theta}{c_h^*}} = 0$$

$$\Delta_{y_2(l)} = 1 - \frac{\theta y_l^*}{\frac{1}{\theta c_l^*}} = 1 - \frac{\theta^4(1-2\gamma)}{\theta^4 - 2\gamma} > 0$$

Example: Implementation

- Note that we cannot interpret wedges as optimal taxes!
- Many different tax systems can implement the optimal allocation
- For example imposing 100% tax if allocations are not optimal (direct mechanism)
- Consider linear income and saving taxes together with lump sum transfers

Example: Implementation

The individual maximization problem is:

$$\max \log(c_1) - \frac{y_1^2}{2} + \Big(\frac{\log(c_h) - \frac{1}{2}(\frac{y_h}{\theta})^2}{2} + \frac{\log(c_l) - \frac{1}{2}(\theta y_l)^2}{2}\Big)$$

subject to

$$\begin{aligned} c_1 &\leq y_1(1-\tau_1) - s + T \quad (\lambda_1) \\ c_h &\leq y_h(1-\tau_h) + s(1-\tau_{s,h}) + T_h \quad (\lambda_h) \\ c_l &\leq y_l(1-\tau_l) + s(1-\tau_{s,l}) + T_l \quad (\lambda_l) \end{aligned}$$

Example: Implementation

The FOCs from the individual maximization problem are:

$$
\begin{aligned}
c_1 &: \frac{1}{c_1} - \lambda_1 = 0 \\
c_h &: \frac{1}{2}\frac{1}{c_h} - \lambda_h \\
c_l &: \frac{1}{2}\frac{1}{c_l} - \lambda_l \\
y_1 &: -y_1 + \lambda_1(1 - \tau_1) = 0 \\
y_h &: -\frac{1}{2}\frac{y_h}{\theta^2} + \lambda_h(1 - \tau_h) \\
y_l &: -\frac{1}{2}\theta^2 y_l + \lambda_l(1 - \tau_l) \\
s &: -\lambda_1 + (1 - \tau_{s,h})\lambda_h + (1 - \tau_{s,l})\lambda_l
\end{aligned}
$$

Example: Implementation

- We obtain the following optimal conditions

$$\frac{1}{c_1} = (1-\tau_{s,h})\frac{1}{2c_h} + (1-\tau_{s,l})\frac{1}{2c_l} \Rightarrow \tau_{s,h} = -2\gamma, \tau_{s,l} = 2\gamma$$

- The condition is satisfied if $\frac{c_h^*}{c_1^*} = 1 + 2\gamma$ and $\frac{c_l^*}{c_1^*} = 1 - 2\gamma$
- Expected revenue from taxes on savings is zero, but these taxes reduce savings because of the generated risk premium
- Furthermore, the other taxes can be set as follows:

$$\begin{aligned} y_1^* c_1^* &= 1 - \tau_1 \Rightarrow \tau_1 = 0 \\ y_h^* c_h^* &= \theta^2(1-\tau_h) \Rightarrow \tau_h = 0 \\ y_l^* c_l^* &= \frac{1-\tau_l}{\theta^2} \Rightarrow \tau_l = 2\gamma\frac{\theta^4 - 1}{\theta^4 - 2\gamma} > 0 \end{aligned}$$

- Transfers are finally found from the budget, $T_l > T_h$

Example: Implementation

- There can be many different implementation: Use means-tested system rather than taxes on capital, that is, receive transfers if saving sufficiently small, otherwise be taxed (Golosov and Tsyvinski, 2006)

Extended Framework

- Non separable consumption and leisure
- Aggregate shocks (Kocherlakota, 2005; Werning, 2007)
- Capital not observable (Abraham and Pavoni, 2003)
- Individual borrowing and lending not observable (Golosov and Tsvinsky, 2006)
- Estate taxation in dynastic model (Fahri and Werning, 2007)
- Gvt without commitment (Revelation Principle does not hold) (Sleet and Yeltekin, 2005; Acemoglu, Golosov and Tsvinsky, 2006)
- On quantification of welfare gain of optimal capital taxes (Fahri and Werning, 2012)