Notes on Mirrlees Taxation (very rough)

Jonathan Heathcote

Static problem:
Agents differ by productivity θ
I values for productivity $\theta_1, ..., \theta_I$
Fraction π_i of each type
Preferences

$$U_i = u(c_i) - v\left(\frac{y_i}{\theta_i}\right)$$

Planner must raise revenue to finance G
Planner puts weight W_i on type i s.t. $\sum_i W_i \pi_i = 1$
An allocation is a vector $\{(c_i, y_i)\}_{i=1}^{I}$
Social welfare is given by

$$\sum_i W_i \pi_i \left\{ u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \right\}$$

Planner can observe y, but not θ
So taxes must be a function of y
Planner's problem is therefore to choose a tax function $T(y)$ such that when agents take this schedule as given and solve

$$\max_{\{c_i, y_i\}} \left\{ u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \right\}$$

$$\text{s.t. } c_i = y_i - T(y_i)$$

the resulting allocations maximize social welfare.

Aside: note that could define $h_i = \frac{y_i}{\theta_i}$ and equivalently write agent's problem as

$$\max_{\{c_i, y_i\}} \left\{ u(c_i) - v\left(h_i\right) \right\}$$

$$\text{s.t. } c_i = \theta_i h_i - T(\theta_i h_i)$$

Suppose $v(x) = (1+\sigma)^{-1} x^{1+\sigma}$. Formulation 1 could be interpred as model in which all agents have same wage, but different disutilities of work. Formulation 2 looks like a model with common preferences but different productivities. But these two models are obviously identical => distinction between preferences and productivity is artificial.

Now the problem is that the optimal T could be a very complicated non-parametric function. How are we supposed to solve for it?

Mirrlees' clever idea.

Instead of thinking of planner picking T think of planner picking allocations directly.

In particular think of planner as offering a menu of different choices $\{(c_i, y_i)\}$ with one pair in this menu intended for each type. The planner can say:

"If you produce income y_i (which I can observe) then you must pay a tax $y_i - c_i$."

But the planner cannot force agents to choose the pair intended for their type, because type is not observed

Thus the planner must incentivize choosing the appropriate allocation by making sure that each type weakly prefers to pick their intended allocation

Thus the Mirrlees problem is

$$\begin{aligned}
& \max_{\{c_i, y_i\}} \sum_i W_i \pi_i \left\{ u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \right\} \\
& s.t. \\
u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \quad & \geq \quad u(c_j) - v\left(\frac{y_j}{\theta_i}\right) \text{ for all } i, j \\
\sum_i \pi_i c_i + G \quad & = \quad \sum_i \pi_i y_i
\end{aligned}$$

There are lots of incentive constraints!

Fortunately most of them will not be binding

Suppose planner wants to redistribute downwards => lower taxes on less productive agents => possible incentive to pretend to be less productive, but no incentive to pretend to be more productive => only downward IC constraints will bind

In fact, only local downward will bind.

Fortunately most of them will not be binding

Suppose planner wants to redistribute downwards => lower taxes on less productive agents => possible incentive to pretend to be less productive, but no incentive to pretend to be more productive => only downward IC constraints will bind

In fact, only local downward will bind. So can simplify problem

$$\begin{aligned}
& \max_{\{c_i, y_i\}} \sum_i W_i \pi_i \left\{ u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \right\} \\
\pi_i \mu_i \quad & : \quad u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \geq u(c_{i-1}) - v\left(\frac{y_{-1}}{\theta_i}\right) \text{ for } i = 2, ..., I \\
\lambda \quad & : \quad \sum_i \pi_i c_i + G \leq \sum_i \pi_i y_i
\end{aligned}$$

$2I + I - 1 + 1$ unknowns

$2I$ FOCs + I constraints

This problem will have a solution. How can we decentralize it? We need to come up with a tax system such that taxes depend on earnings, and all agents are on their FOC

$$u'(c_i)\theta_i(1-T'(y_i)) = v'\left(\frac{y_i}{\theta_i}\right)$$

and

$$c_i = y_i - T(y_i)$$

Note that marginal and average tax rates are only exactly pinned down at grid points.

FOCs

$$\begin{aligned}
c_i &: \quad W_i\pi_i u'(c_i) + \pi_i\mu_i u'(c_i) - \pi_{i+1}\mu_{i+1}u'(c_i) - \lambda\pi_i = 0 \\
y_i &: \quad -W_i\pi_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} - \pi_i\mu_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} + \pi_{i+1}\mu_{i+1}v'\left(\frac{y}{\theta_{i+1}}\right)\frac{1}{\theta_{i+1}} + \lambda\pi_i = 0
\end{aligned}$$

Note that for $i = I$ the $i+1$ terms are absent, so can combine the two to give

$$\begin{aligned}
W_i\pi_i u'(c_i) + \pi_i\mu_i u'(c_i) &= W_i\pi_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} + \pi_i\mu_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} \\
\left(W_i\pi_i + \pi_i\mu_i\right)\theta_i u'(c_i) &= \left(W_i\pi_i + \pi_i\mu_i\right)v'\left(\frac{y_i}{\theta_i}\right)
\end{aligned}$$

So no distortion at the top.

More generically, suppose $v(x) = (1+\sigma)^{-1}x^{1+\sigma}$, so $v'(x) = x^{\sigma}$

$$v'\left(\frac{y_i}{\theta_{i+1}}\right) = v'\left(\frac{y_i}{\theta_i}\frac{\theta_i}{\theta_{i+1}}\right) = v'\left(\frac{y_i}{\theta_i}\right)\left(\frac{\theta_i}{\theta_{i+1}}\right)^{\sigma}$$

$$\begin{aligned}
\left(W_i\pi_i + \pi_i\mu_i - \pi_{i+1}\mu_{i+1}\right)u'(c_i) - \lambda\pi_i &= 0 \\
\left(-W_i\pi_i - \pi_i\mu_i + \pi_{i+1}\mu_{i+1}\left(\frac{\theta_i}{\theta_{i+1}}\right)^{1+\sigma}\right)\frac{1}{\theta_i}v'\left(\frac{y_i}{\theta_i}\right) + \lambda\pi_i &= 0
\end{aligned}$$

So we have a "wedge" in the intra-temporal FOC given by

$$wedge = (1 - tax) = \frac{\left(W_i\pi_i + \pi_i\mu_i - \pi_{i+1}\mu_{i+1}\right)}{\left(W_i\pi_i + \pi_i\mu_i - \pi_{i+1}\mu_{i+1}\left(\frac{\theta_i}{\theta_{i+1}}\right)^{1+\sigma}\right)}$$

Note that wedge is less than one, so tax is positive => labor supply is distorted => can't achieve first best

What about at the bottom. There is no IC for $i = 1$, equivalently $\mu_1 = 0$.

How do we solve this numerically?
Guess λ
Guess c_1
Solve for μ_2 for FOC for c_1
Solve for y_1 for FOC for y_1
3 equations (2 FOCs and IC_2) to solve for c_2, y_2 and μ_3
Iterate upwards through the grid
At $I-1$ we solve for μ_I
Then we have 2 FOCs at I to solve for c_I and y_I
Check IC_I and adjust c_1 if not satisfied
Finally check resource constraint and adjust λ