Recitations 26

1 Static Mirrlees taxation

Standard assumption in the Ramsey literature is that lump sum taxes are not allowed. Why aren't lump sum taxes used in practice?
One reason for this is they require truthful elicitation of agents characteristics, which might not be publicly observable. Moreover, agents might not have an incentive to reveal these characteristics truthfully.

We will consider a mechanism design problem in which agents true ability types are private and allow the designer to use arbitrary mechanisms and transfer schedules to achieve efficiency. Next, will consider implementations/decentralizations.

1.1 A Two Type Example

Consider an environment with a continuum of HHs characterized by a productivity level $\theta \in \Theta = \{\theta_H, \theta_L\}$ with $\theta_H > \theta_L > 0$.
A household of type θ who works l hours can produce $y = \theta l$ of output. Let $\pi(\theta)$ denote the probability that given household is of type θ. By the LLN, this is also the fraction of HHs with productivity θ.

Household preferences are given by $u(c) - v(l)$ but we will use $l = \frac{y}{\theta}$ to define the preferences as $U(c, y, \theta) = u(\mathrm{c}) - v\left(\frac{y}{\theta}\right)$. Assume $u' > 0 < u''$ and $v', v'' > 0$.

Suppose first that HH productivities are public information. Under full information a utilitarian planner (cares about all types equally) solves

$$\max_{c(\theta), y(\theta)} \pi\left(\theta_H\right)\left[u\left(c\left(\theta_H\right)\right) - v\left(\frac{y\left(\theta_H\right)}{\theta_H}\right)\right] + \pi\left(\theta_L\right)\left[u\left(c\left(\theta_L\right)\right) - v\left(\frac{y\left(\theta_L\right)}{\theta_L}\right)\right]$$

subject to

$$\pi\left(\theta_{\mathrm{H}}\right) \mathrm{c}\left(\theta_{\mathrm{H}}\right) + \pi\left(\theta_{\mathrm{L}}\right) c\left(\theta_{\mathrm{L}}\right) \leqslant \pi\left(\theta_{\mathrm{H}}\right) \mathrm{y}\left(\theta_{\mathrm{H}}\right) + \pi\left(\theta_{\mathrm{L}}\right) \mathrm{y}\left(\theta_{\mathrm{L}}\right)$$

Let μ be the multiplier on the RC. Then the the FOCs are

$$u'(c(\theta_H)) = \mu = u'(c(\theta_L))$$
$$\frac{1}{\theta_H} v'(l(\theta_H)) = \mu = \frac{1}{\theta_L} v'(l(\theta_L))$$

Which implies

$$c(\theta_H) = c(\theta_L)$$
$$c(\theta) = \tfrac{1}{\theta} v'(l(\theta))$$
$$\tfrac{v'(l(\theta_H))}{v'(l(\theta_L))} = \tfrac{\theta_H}{\theta_L} > 1$$

where the last equation implies that $l(\theta_H) > l(\theta_L)$ since v is convex.

Now suppose that θ is private information. It is easy to see that the above allocation is not incentive compatible. A high type households strictly prefers to pretend to be a low type since the consumption levels are the same but hours worked is lower. From the revelation principle, that we can restrict ourselves to direct revelation mechanisms.

Definition 1.1 (Direct revelation mechanism)**.** *consists of action/message sets* $M_i, i \in [0,1]$ *such that for each* $i, M_i = \Theta_i$ *and outcome functions* (c, y) *where* $c, y : \Theta \to \mathbb{R}_+$.

Since there is no aggregate uncertainty (LLN), we will consider mechanisms that treat households anonymously, i.e. mechanisms that are independent of i.

Definition 1.2 (Revelation mechanism)**.** *is*

1. ***Incentive compatible (IC)*** *if and only if*

$$u(c_H) - v(l_H) \geq u(c_L) - v\left(\frac{\theta_L}{\theta_H} l_L\right) \tag{1.1}$$

$$u(c_L) - v(l_L) \geq u(c_H) - v\left(\frac{\theta_H}{\theta_L} l_H\right)$$

2. ***Resource feasible (FEAS)*** *if and only if*

$$\pi(\theta_H) c(\theta_H) + \pi(\theta_L) c(\theta_L) \leqslant \pi(\theta_H) y(\theta_H) + \pi(\theta_L) y(\theta_L) \tag{1.2}$$

Then, the Planner's/Mechanism designer's problem is

$$\max_{c(\theta), y(\theta)} \pi(\theta_H) \left[u(c(\theta_H)) - v\left(\frac{y(\theta_H)}{\theta_H}\right)\right] + \pi(\theta_L) \left[u(c(\theta_L)) - v\left(\frac{y(\theta_L)}{\theta_L}\right)\right]$$

s.t. IC1, IC2, FEAS

Notice that there are two incentive compatibility constraints, one for the high and one for the low type.

1.2 The relaxed problem

This problem is typically not concave/convex - concave objective function and a convex constraint set. This is because the $u(c)'s$ (and $v's$) appear on both the left and right hand side of the (IC) constraints.

We need to reformulate our problem so we can use KKT.

Before that let's note some simple properties of contracts, i.e. combinations of (c_L, y_L) and (c_H, y_H) that must be true if (FEAS), (IC1), and (IC2) are all satisfied.

- Suppose $c_H > c_L$ but $y_H \leq y_L$. If this were true, then,

$$u(c_H) - v\left(\frac{y_H}{\theta_L}\right) > u(c_L) - v\left(\frac{y_H}{\theta_L}\right)$$

  since $c_H > c_L$ and $u(\cdot)$ is monotone. Moreover,

$$u(c_L) - v\left(\frac{y_H}{\theta_L}\right) \geq u(c_L) - v\left(\frac{y_L}{\theta_L}\right)$$

  since $y_L \geq y_H$ and $v(\cdot)$ is monotone. Thus,

$$u(c_H) - v\left(\frac{y_H}{\theta_L}\right) > u(c_L) - v\left(\frac{y_L}{\theta_L}\right)$$

  But this violates (IC2) and hence, these types of allocations are not feasible.

- A similar argument shows that combinations with $c_H \geq c_L$ and $y_H < y_L$ also are not feasible.

- Suppose $c_H < c_L$ but $y_L \leq y_H$. If this were true, as above, we would have

$$u(c_L) - v\left(\frac{y_L}{\theta_H}\right) > u(c_H) - v\left(\frac{y_H}{\theta_H}\right)$$

  i.e. (IC1) would be violated.

- A similar argument holds if $c_H \leq c_L$ but $y_L < y_H$. So, this cannot be the case in the solution.

We can summarize these in the following lemma.

Lemma 1.3. *If the contract (c_L, y_L) and (c_H, y_H) satisfies* (FEAS), (IC1), *and* (IC2) *in Problem (SP2), then, one of the following three configurations must hold:*

1. $c_H > c_L$ *and* $y_H > y_L$
2. $c_L > c_H$ *and* $y_L > y_H$; *or,*
3. $c_L = c_H$ *and* $y_L = y_H$

Here we show that only first allocation can be optimal and other two are impossible. To do so we will use **Variational methods**

Proof. • $c_L > c_H$ and $y_L > y_H$; is not possible.

Assume $(c_L, y_L) > (c_H, y_H)$. By (IC2),

$$u(c_L) - u(c_H) - \left[v\left(\frac{y_L}{\theta_L}\right) - v\left(\frac{y_H}{\theta_L}\right)\right] \geq 0$$

then by taking integral representation of last two we obtain:

$$u(c_L) - u(c_H) \geq \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy$$

Notice that the first term, $u(c_L) - u(c_H)$ is positive, since c_L is assumed to be larger than c_H. Also, since $v'(\cdot) > 0$ and $y_L > y_H$, it follows that the second term is also positive. But since $\theta_H > \theta_L$ and $v(\cdot)$ is convex, it follows that $v'(y/\theta_H) < v'(y/\theta_L)$ for all y, and, as a result,

$$\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy < \int_{y_H}^{y_L} v\left(\frac{y}{\theta_L}\right) dy$$

Since $\theta_H > \theta_L$, we also have:

$$\frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy < \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy$$

Thus,

$$\begin{aligned} u(c_L) - u(c_H) &\geq \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy \\ &> \frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy \end{aligned}$$

implying that,

$$u(c_L) - u(c_H) - \frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy > 0$$

$$\therefore u(c_L) - u(c_H) - \left[v\left(\frac{y_L}{\theta_H}\right) - v\left(\frac{y_H}{\theta_H}\right)\right] > 0$$

That is high-type agents prefer (c_L, y_L) over (c_H, y_H), violating (IC1).

- $c_L = c_H$ and $y_L = y_H$

To show this formally, suppose (c, y) denotes the common consumption/production pair. We consider the following cases:

1. If

$$u'(c) < \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right)$$

since $\theta_H > \theta_L$, we have:

$$u'(c) < \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right)$$

Therefore, decreasing c and y at the same time by a small amount will keep the (IC)s holding and will be strictly better for both types.

2. If

$$u'(c) = \frac{1}{\theta_H} v' \left(\frac{y}{\theta_H} \right)$$

again, we have:

$$u'(c) < \frac{1}{\theta_L} v' \left(\frac{y}{\theta_L} \right)$$

Hence, decreasing consumption and production of the low types would make them better off, while high types have no incentives to deviate to the new allocation.

3. If

$$u'(c) > \frac{1}{\theta_H} v' \left(\frac{y}{\theta_H} \right) \text{ and } u'(c) > \frac{1}{\theta_L} v' \left(\frac{y}{\theta_L} \right)$$

by the same argument as in the first case, it is optimal to increase y and c at the same time. The case of

$$u'(c) > \frac{1}{\theta_H} v' \left(\frac{y}{\theta_H} \right) \text{ and } u'(c) = \frac{1}{\theta_L} v' \left(\frac{y}{\theta_L} \right)$$

is not optimal either, by the same logic as in the second case.

4. If

$$u'(c) > \frac{1}{\theta_H} v' \left(\frac{y}{\theta_H} \right) \text{ and } u'(c) < \frac{1}{\theta_L} v' \left(\frac{y}{\theta_L} \right)$$

an increase in the consumption of high types followed by an increase in their production, and a decrease in the consumption of low types followed by a decrease in their production would leave both types better off.

□

Lemma 1.4. *(IC1) holds with equality*

Proof. We will use **perturbation argument**.We will do this by supposing it is false and constructing a better contract. The dominating contract that we will construct will have better insurance over c without disrupting (IC1). Suppose that (IC1) is not satisfied at equality;

$$u(c_H) - v\left(\frac{y_H}{\theta_H}\right) > u(c_L) - v\left(\frac{y_L}{\theta_H}\right)$$

Notice that, if this holds, by continuity, it will still hold if we add a bit to c_L and subtract a bit from c_H

$$u(c_H - \varepsilon) - v\left(\frac{y_H}{\theta_H}\right) > u(c_L + \delta) - v\left(\frac{y_L}{\theta_H}\right)$$

as long as ε and δ are small enough. Consider the alternative contract given by $(c_H - \varepsilon, y_H)$ and $(c_L + \delta, y_L)$. Choose $\delta = \pi_H \varepsilon / \pi_L$. Then, if ε is small enough, (IC1) will still hold, and (FEAS) becomes:

$$\begin{aligned}\pi_H (c_H - \varepsilon) + \pi_L (c_L + \delta) &= \pi_H c_H + \pi_L c_L + \pi_H \varepsilon - \pi_L \frac{\pi_H}{\pi_L} \varepsilon \\ &= \pi_H c_H + \pi_L c_L\end{aligned}$$

Thus, (FEAS) will hold because we didn't change y_L or y_H, and because of the way we constructed δ.

So, we only need to show that welfare goes up from this change, even when ε is small but positive. To see this note that the change in welfare is given by:

$$\Delta W = \pi_H \left[u (c_H - \varepsilon) - u (c_H)\right] + \pi_L \left[u \left(c_L + \frac{\pi_H}{\pi_L} \varepsilon\right) - u (c_L)\right]$$

The terms involving the y 's do not appear in this, since they are unchanged. If we take the derivative with respect to ε, at $\varepsilon = 0$, we have:

$$\begin{aligned}\left.\frac{d\Delta W}{d\varepsilon}\right|_{\varepsilon=0} &= -\pi_H u' (c_H) + \pi_L \frac{\pi_H}{\pi_L} u' (c_L) \\ &= -\pi_H u' (c_H) + \pi_H u' (c_L) \\ &= \pi_H \left[u' (c_L) - u' (c_H)\right] \\ &> 0\end{aligned}$$

since $c_L < c_H$, and $u(\cdot)$ is assumed to be strictly concave. □

Lemma 1.5. *If (IC1) holds with equality then (IC2) is satisfied for* $c_H > c_L$ *and* $y_H > y_L$

Proof. Suppose not and that this constraint was violated. Then

$$\begin{aligned}& u (c (\theta_L)) - v \left(\frac{y (\theta_L)}{\theta_L}\right) < u (c (\theta_H)) - v \left(\frac{y (\theta_H)}{\theta_L}\right) \\ \Longrightarrow & v \left(\frac{y (\theta_H)}{\theta_L}\right) - v \left(\frac{y (\theta_L)}{\theta_L}\right) < u (c (\theta_H)) - u (c (\theta_L)) \\ \Longrightarrow & \frac{1}{\theta_L} \int_{y(\theta_L)}^{y(\theta_H)} v' \left(\frac{y}{\theta_L}\right) dy < u (c (\theta_H)) - u (c (\theta_L)) \\ \Longrightarrow & \frac{1}{\theta_H} \int_{y(\theta_L)}^{y(\theta_H)} v' \left(\frac{y}{\theta_H}\right) dy < \frac{1}{\theta_L} \int_{y(\theta_L)}^{y(\theta_H)} v' \left(\frac{y}{\theta_L}\right) dy < u (c (\theta_H)) - u (c (\theta_L)) \\ \Longrightarrow & v \left(\frac{y (\theta_H)}{\theta_H}\right) - v \left(\frac{y (\theta_L)}{\theta_H}\right) < u (c (\theta_H)) - u (c (\theta_L)) \\ \Longrightarrow & u (c (\theta_L)) - v \left(\frac{y (\theta_L)}{\theta_H}\right) < u (c (\theta_H)) - v \left(\frac{y (\theta_H)}{\theta_H}\right)\end{aligned}$$

which contradicts the IC for the high type holding with equality. □

Finally we are ready to use KKT to solve **Relaxed problem**

$$\max_{c(\theta),y(\theta)} \pi(\theta_H)\left[u(c(\theta_H)) - v\left(\frac{y(\theta_H)}{\theta_H}\right)\right] + \pi(\theta_L)\left[u(c(\theta_L)) - v\left(\frac{y(\theta_L)}{\theta_L}\right)\right]$$

subject to

$$u(c(\theta_H)) - v\left(\tfrac{y(\theta_H)}{\theta_H}\right) = u(c(\theta_L)) - v\left(\tfrac{y(\theta_L)}{\theta_H}\right)$$
$$\pi(\theta_H)c(\theta_H) + \pi(\theta_L)c(\theta_L) \leqslant \pi(\theta_H)y(\theta_H) + \pi(\theta_L)c(\theta_L)$$

Let λ be the multiplier on the first constraint and μ on the second.

The FOCs are

$$\pi(\theta_\mathrm{H})u'(c(\theta_\mathrm{H})) + \lambda u'(c(\theta_\mathrm{H})) - \pi(\theta_\mathrm{H})\mu = 0 \tag{1.3}$$

$$\pi(\theta_\mathrm{L})u'(c(\theta_\mathrm{L})) - \lambda u'(c(\theta_\mathrm{L})) - \pi(\theta_\mathrm{L})\mu = 0 \tag{1.4}$$

$$-\frac{\pi(\theta_\mathrm{H})}{\theta_\mathrm{H}}v'\left(\frac{\mathrm{y}(\theta_\mathrm{H})}{\theta_\mathrm{H}}\right) - \lambda\frac{1}{\theta_\mathrm{H}}v'\left(\frac{\mathrm{y}(\theta_\mathrm{H})}{\theta_\mathrm{H}}\right) + \pi(\theta_\mathrm{H})\mu = 0 \tag{1.5}$$

$$-\frac{\pi(\theta_\mathrm{L})}{\theta_\mathrm{L}}v'\left(\frac{\mathrm{y}(\theta_\mathrm{L})}{\theta_\mathrm{L}}\right) + \lambda\frac{1}{\theta_\mathrm{H}}v'\left(\frac{\mathrm{y}(\theta_\mathrm{L})}{\theta_\mathrm{H}}\right) + \pi(\theta_\mathrm{L})\mu = 0 \tag{1.6}$$

Combining (3) and 5) we obtain

$$u'(c(\theta_H)) = \frac{1}{\theta_H}\nu'\left(\frac{y(\theta_H)}{\theta_H}\right)$$

This says that the just as in the unconstrained problem, for the high type, the marginal utility of c equals the marginal disutility of working. In particular, the allocation for the high type households is ex-post efficient. This is sometimes referred to as "**no-distortion at the top**".

The mechanical reason for this is that, no type wants to pretend to be the high type and thus the planner does not need to distort. This will not be true for the low type. Next, combine (3) and (4) :

$$\frac{u'(c(\theta_H))}{u'(c(\theta_L))} = \frac{\pi(\theta_H)}{\pi(\theta_L)}\frac{[\pi(\theta_L) - \lambda]}{[\pi(\theta_H) + \lambda]} = \frac{\pi(\theta_H)\pi(\theta_L) - \lambda\pi(\theta_H)}{\pi(\theta_H)\pi(\theta_L) + \lambda\pi(\theta_L)} < 1$$
$$\Rightarrow u'(c(\theta_H)) < u'(c(\theta_L)) \quad \Rightarrow c(\theta_H) > c(\theta_L)$$

But then given the IC holds with equality, it must be that

$$v\left(\frac{\mathrm{y}(\theta_\mathrm{H})}{\theta_\mathrm{H}}\right) > v\left(\frac{\mathrm{y}(\theta_\mathrm{L})}{\theta_\mathrm{H}}\right)$$
$$\Rightarrow \mathrm{y}(\theta_\mathrm{H}) > \mathrm{y}(\theta_\mathrm{L})$$

To see the allocation for the low type is distorted, combine (4) and (6) to obtain

$$u'(c(\theta_L)) = \frac{1}{\theta_L}v'\left(\frac{y(\theta_L)}{\theta_L}\right) + \frac{\lambda}{\pi(\theta_L)}\left[u'(c(\theta_L)) - \frac{1}{\theta_H}v'\left(\frac{y(\theta_L)}{\theta_H}\right)\right] >$$
$$> \frac{1}{\theta_L}v'\left(\frac{y(\theta_L)}{\theta_L}\right) + \frac{\lambda}{\pi(\theta_L)}\left[u'(c(\theta_H)) - \frac{1}{\theta_H}v'\left(\frac{y(\theta_H)}{\theta_H}\right)\right]$$
$$= \frac{1}{\theta_L}v'\left(\frac{y(\theta_L)}{\theta_L}\right)$$

To make sure that indeed we have solution which maximizes welfare need to check SOCs

$$\frac{d^2L}{d(c_H)^2}: \quad \pi(\theta_{\mathrm{H}})\,u''(c(\theta_{\mathrm{H}})) + \lambda u''(c(\theta_{\mathrm{H}}))$$

$$\frac{d^2L}{d(c_L)^2}: \quad \pi(\theta_{\mathrm{L}})\,u''(c(\theta_{\mathrm{L}})) - \lambda u''(c(\theta_{\mathrm{L}}))$$

$$\frac{d^2L}{d(y_H)^2}: \quad -\frac{\pi(\theta_{\mathrm{H}})}{\theta_{\mathrm{H}}^2}v''\left(\frac{\mathrm{y}(\theta_{\mathrm{H}})}{\theta_{\mathrm{H}}}\right) - \lambda\frac{1}{\theta_{\mathrm{H}}^2}v''\left(\frac{\mathrm{y}(\theta_{\mathrm{H}})}{\theta_{\mathrm{H}}}\right)$$

$$\frac{d^2L}{d(y_L)^2}: \quad -\frac{\pi(\theta_{\mathrm{L}})}{\theta_{\mathrm{L}}^2}v''\left(\frac{\mathrm{y}(\theta_{\mathrm{L}})}{\theta_{\mathrm{L}}}\right) + \lambda\frac{1}{\theta_{\mathrm{H}}}v''\left(\frac{\mathrm{y}(\theta_{\mathrm{L}})}{\theta_{\mathrm{H}}^2}\right)$$

And all other cross derivatives are 0. Observe that from (1)

$$-\pi(\theta_H) < \lambda = \frac{\pi(\theta_L)u'(c_L) - \pi(\theta_H)u'(c_H)}{u'(c_H) + u'(c_L)} < \pi_(\theta_L)$$

Then all second order non zero derivatives are negative, so Hessian of our Langrasian is negative definite on whole $\mathbb{R}^4$ (in particular it is negative definite on kernel of linear epimorphism of Jacobian generated by constraints). So indeed we have solution to our problem whih maximizes welfare.

1.3 Don't distort at the top

Above tells us nothing about implementation, i.e. whether there exist tax systems, for example such that the equilibrium, given the tax system gives efficient allocation. We turn to this next. In particular, we will show that a non-linear income tax schedule can implement the efficient allocation.

Denote the optimal mechanism by (c^*, y^*). Define a tax function $T(y) = y - c$ if $y \in \{y^*(\theta_H), y^*(\theta_L)\}$ and $T(y) = y$ otherwise. Given this tax function, the household of type θ solves:

$$\max_{c,y} u(c) - v\left(\frac{y}{\theta}\right)$$

$$\text{st} \quad c \leq y - T(y)$$

The first order condition is

$$u'(c)\left(1-T'(y)\right)=\frac{1}{\theta}v'\left(\frac{y}{\theta}\right)$$

Comparing this equation to the one in the planning problem implies that $T'\left(y_H^*\right)=0$ and $T'\left(y_L^*\right)>0$

1.4 Mirrlees with a continuum of types

We now consider a problem with a continuum of types. We characterize the efficient allocation and derive the Diamond-Mirrlees-Saez formula.

The main issue for continuum of types is how to simplify IC. In the two type case, we just dropped one but here we can not do things like that. But we have already seen how to deal with these constraints! We will use similar Myerson(1981)- like techniques and replace incentive compatibility with an local condition and a monotonicity condition.

As before, an allocation (c, y) is incentive compatible (GIC) if and only if

$$U(\theta)\equiv u(c(\theta))-v\left(\frac{y(\theta)}{\theta}\right)\geq u(c(\hat{\theta}))-v\left(\frac{y(\hat{\theta})}{\theta}\right)\equiv u(\hat{\theta},\theta)\quad\forall\theta,\theta\in\Theta$$

Lemma 1.6. *An allocation* (c,y) *satisfies global incentive compatibility* $\Longleftrightarrow$

1. $\mathrm{y}(\theta)$ *is increasing in* θ
2. $$u'(c(\theta))c'(\theta)=\frac{1}{\theta}y'(\theta)v'\left(\frac{y(\theta)}{\theta}\right)$$

As a result of lemma, we can now write down a relaxed planning problem

$$\max_{c,y}\int W\left(u(c(\theta))-v\left(\frac{y(\theta)}{\theta}\right)\right)dF$$

subject to

$$\begin{gathered}\int_\Theta \mathrm{c}(\theta)dF(\theta)\leqslant\int_\Theta \mathrm{y}(\theta)dF(\theta)\\ \mathrm{u}'(\mathrm{c}(\theta))\mathrm{c}'(\theta)=\tfrac{1}{\theta}y'(\theta)v'\left(\tfrac{\mathrm{y}(\theta)}{\theta}\right)\\ \mathrm{y}'(\theta)\geqslant 0\end{gathered}$$

Here F is the cdf of θ and W is a general weighting function instead of just assuming a utilitarian planner. This problem is still pretty intractable. We will use one more trick: replace the derivative condition with an **envelope condition**. In particular the derivative condition is equivalent to

$$\mathrm{U}(\theta)=\max_{\hat{\theta}\in\Theta}\mathrm{U}(\hat{\theta},\theta)$$

The envelope condition for this maximization problem is

$$\begin{aligned} \mathrm{u}'(\theta) &= \frac{\partial}{\partial\theta} u(c(\hat{\theta})) - v\left(\frac{\mathrm{y}(\hat{\theta})}{\theta}\right)\Bigg|_{\hat{\theta}=\theta} \\ &= \frac{\mathrm{y}(\theta)}{\theta^2} v'\left(\frac{\mathrm{y}(\theta)}{\theta}\right) \end{aligned}$$

For a final time, the planning problem is

$$\max_{\mathrm{c},y} \int W(\mathrm{U}(\theta)) dF(\theta)$$

$$\begin{gathered} \text{st} \quad \mathrm{U}(\theta) = u(c(\theta)) - v\left(\tfrac{\mathrm{y}(\theta)}{\theta}\right) \\ \textstyle\int_\Theta \mathrm{c}(\theta)\mathrm{dF}(\theta) \leqslant \int_\Theta \mathrm{y}(\theta)\mathrm{dF}(\theta) \\ \mathrm{U}'(\theta) = \tfrac{\mathrm{y}(\theta)}{\theta^2} v'\left(\tfrac{\mathrm{y}(\theta)}{\theta}\right) \\ \mathrm{y}'(\theta) \geqslant 0 \end{gathered}$$

A common mechanism design trick is to drop the monotonicity condition and check that the result allocation satisfies it ex-post. If it is violated then it usually means that there is "bunching". To deal with this situation we use an **ironing method**- see Myerson.

To solve the problem above we use techniques from the **calculus of variation**. We can write the lagrangian

$$\begin{aligned} \mathcal{L} = &\int W(\mathrm{U}(\theta)) dF(\theta) + \int \gamma(\theta) \left[\mathrm{u}(\mathrm{c}(\theta)) - v\left(\frac{\mathrm{y}(\theta)}{\theta}\right) - \mathrm{U}(\theta)\right] \mathrm{d}\theta \\ &+ \lambda \left[\int_\Theta \mathrm{y}(\theta)\mathrm{dF}(\theta) - \int_\Theta \mathrm{c}(\theta)\mathrm{dF}(\theta)\right] \\ &+ \int \mu(\theta) \left[\mathrm{U}'(\theta) - \frac{\mathrm{y}(\theta)}{\theta^2} v'\left(\frac{\mathrm{y}(\theta)}{\theta}\right)\right] \mathrm{d}\theta \end{aligned}$$

Lets first deal with the $\mathrm{U}'(\theta)$ term:

$$\begin{aligned} \int \mu(\theta)\mathrm{U}'(\theta)\mathrm{d}\theta &= \int \mu(\theta)\mathrm{dU}(\theta) \\ &= \mu(\bar{\theta})\mathrm{U}(\bar{\theta}) - \mu(\underline{\theta})\mathrm{U}(\underline{\theta}) - \int \mu'(\theta)\mathrm{U}(\theta)\mathrm{d}\theta \end{aligned}$$

Now substitute this back in

$$\mathcal{L} = \int W(\mathrm{U}(\theta)) dF(\theta) + \int \gamma(\theta) \left[\mathrm{u}(\mathrm{c}(\theta)) - v\left(\frac{\mathrm{y}(\theta)}{\theta}\right) - \mathrm{U}(\theta)\right] \mathrm{d}\theta$$

$$+\lambda\left[\int_\Theta y(\theta)dF(\theta)-\int_\Theta c(\theta)dF(\theta)\right]$$
$$-\int \mu(\theta)\frac{y(\theta)}{\theta^2}v'\left(\frac{y(\theta)}{\theta}\right)d\theta-\int \mu'(\theta)U(\theta)d\theta+\mu(\bar\theta)u(\bar\theta)-\mu(\underline{\theta})U(\underline{\theta})$$

and take first order conditions:

$$U(\theta): W'(\mathrm{U}(\theta))f(\theta)-\gamma(\theta)-\mu'(\theta)=0 \tag{1.7}$$

$$c(\theta): \gamma(\theta)u'(c(\theta))-\lambda f(\theta)=0 \tag{1.8}$$

$$y(\theta): -\gamma(\theta)\frac{1}{\theta}\nu'\left(\frac{y(\theta)}{\theta}\right)+\lambda f(\theta)-\mu(\theta)\left[\frac{1}{\theta^2}\nu'\left(\frac{y(\theta)}{\theta}\right)+\frac{y(\theta)}{\theta^3}v''\left(\frac{y(\theta)}{\theta}\right)\right]=0 \tag{1.9}$$

We also have two boundary conditions: $\mu(\bar\theta)=\mu(\underline{\theta})=0$.

To see why these must hold notice that if $\mu(\bar\theta)>0(\mu(\underline{\theta})>0)$ then the planner would like to set $\mathrm{U}(\bar\theta)=\infty(\mathrm{U}(\underline{\theta})=-\infty)$ which would clearly violate incentive constraints. Using (7) and (8) we have

$$\begin{aligned}\mu(\theta)&=\int_\theta^{\bar\theta}\left[\gamma(z)-W'(\mathrm{U}(z))f(z)\right]\mathrm{d}z\\&=\int_\theta^{\bar\theta}\left[\frac{\lambda f(z)}{u'(c(z))}-W'(\mathrm{U}(z))f(z)\right]\mathrm{d}z\end{aligned}$$

Then (9) becomes

$$\lambda f(\theta)-\frac{\lambda f(\theta)}{u'(c(\theta))}\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right)=$$
$$=\left[\frac{1}{\theta^2}\nu'\left(\frac{y(\theta)}{\theta}\right)+\frac{y(\theta)}{\theta^3}v''\left(\frac{y(\theta)}{\theta}\right)\right]\int_\theta^{\bar\theta}\left[\frac{\lambda f(z)}{u'(c(z))}-W'(u(z))f(z)\right]dz$$

Divide both sides by $\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right)$

$$\frac{1}{\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right)}-\frac{1}{u'(c(\theta))}=$$
$$=\frac{1-F(\theta)}{\theta f(\theta)}\left[1+\frac{y(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)}\right]\int_\theta^{\bar\theta}\left[\frac{1}{u'(c(z))}-\frac{W'(\mathrm{U}(z))}{\lambda}\right]\frac{dF(z)}{1-F(\theta)}$$

Now recall that we are interested in implementing the efficient allocation with a tax function T(y). In the decentralized problem $u'(\mathrm{c}(\theta))\left(1-\mathrm{T}'(\mathrm{y}(\theta))\right)=\frac{1}{\theta}v'\left(\frac{\mathrm{y}(\theta)}{\theta}\right)$. Therefore the above equation can be written as

$$\frac{1}{u'(c(\theta))\left(1-T'(y(\theta))\right)}-\frac{1}{u'(c(\theta))}=$$

$$=\frac{1-F(\theta)}{\theta f(\theta)}\left[1+\frac{y(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)}\right]\int_{\theta}^{\bar{\theta}}\left[\frac{1}{u'(c(z))}-\frac{W'(\mathrm{U}(z))}{\lambda}\right]\frac{dF(z)}{1-F(\theta)}$$

$$\frac{T'(\mathrm{y})}{1-T'(\mathrm{y})}=u'(c(\theta))\frac{1-\mathrm{F}(\theta)}{\theta f(\theta)}\left[1+\frac{\mathrm{y}(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)}\right]\int_{\theta}^{\bar{\theta}}\left[\frac{1}{u'(c(z))}-\frac{W'(\mathrm{U}(z))}{\lambda}\right]\frac{dF(z)}{1-\mathrm{F}(\theta)} \tag{1.10}$$

Note that the LHS is increasing in τ. This is the famous **Diamond-Mirrlees-Saez formula**. This equation says that the optiaml marginal tax rates are determined by three things:

1. **The hazard rate** $\frac{1-F(\theta)}{f(\theta)}$ (f of the tail of the type distribution) In particular, for bounded distributions marginal taxes should be zero at the top. On the other hand if the distribution has fat tails, like the Pareto distribution then this term is positive.
2. **Labor supply elasticity** $\frac{v'\left(\frac{y(\theta)}{\theta}\right)}{\frac{y(\theta)}{\theta}v''\left(\frac{y(\theta)}{\theta}\right)}$: captures the effect of Frisch elasticity of labor supply (Captures the substitution effect of a marginal change in wage). The formula suggests that if labor is very elastic, then marginal tax rates should be low.
3. **Concern for redistribution**: If the planner loves redistribution then loosely the term $\int_{\theta}^{\bar{\theta}}\frac{W'(\mathrm{U}(z))}{\lambda}$ is small since planner cares more about the lower types. As a result marginal tax rates are higher.