Class 2

Jonathan Heathcote
FRB Minneapolis

Pubic Economics, U of M, Fall 2021

Basic Approaches in Literature

- Ramsey
  - Parametric functional form for taxes (e.g. affine)
  - Solving for optimum schedule just means solving for a few parameter values
- Mirrlees
  - Solve for optimal non-parametric schedule
  - No ad hoc restrictions $\Rightarrow$ should be able to deliver higher welfare
  - Might be harder to compute
- Shape of optimal schedule still hotly debated, e.g.:
  - Should transfers be targeted to the poorest, and rapidly phased out as income rises? (means-tested transfers)
  - Or should transfers be more universal? (UBI)

Mirrlees Taxation

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

$$
\begin{aligned}
U_i &= u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \\
& \log(c_i) - \frac{\left(\frac{y_i}{\theta_i}\right)^{1+\sigma}}{1+\sigma}
\end{aligned}
$$

- Planner must raise revenue to finance G
- Planner puts weight W_i on type i s.t. $\sum_i W_i \pi_i = 1$

Mirrlees Taxation

- 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 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.

Mirrlees Clever Idea

- Optimal T could be a very complicated non-parametric function
- Instead of thinking of planner picking T think of planner picking allocations directly.
- Planner offers a menu of different choices $\{(c_i, y_i)\}$ with one pair in this menu intended for each type
- Planner says: "If you produce income y_i (which I can observe) then you must pay a tax $y_i - c_i$."
- But planner cannot force agents to choose the pair intended for their type, because type is not observed
- Thus planner must incentivize each type to pick their intended allocation

Mirrlees Problem

- Thus the Mirrlees problem is

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

- There are lots of incentive constraints!
- Fortunately most of them will not be binding
- Planner wants to redistribute downwards $\Rightarrow$ only downward IC constraints will bind
- In fact, only local downward constraints will bind.

Simplified Problem

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

$$\pi_i\mu_i \quad : \quad u(c_i) - v\left(\frac{y_i}{\theta_i}\right) \geq u(c_{i-1}) - v\left(\frac{y_{i-1}}{\theta_i}\right) \text{ for } i = 2, ..., I$$

$$\lambda \quad : \quad \sum_i \pi_i c_i + G \leq \sum_i \pi_i y_i$$

- FOCs (recall no IC constraint for $i = 1 \Rightarrow \mu_1 = 0$)

$$c_i \quad : \quad W_i\pi_i u'(c_i) + \pi_i\mu_i u'(c_i) - \lambda\pi_i + \pi_{i+1}\mu_{i+1}u'(c_i) = 0 \text{ for } i = 1, ..., I-1$$

$$y_i \quad : \quad -W_i\pi_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} - \pi_i\mu_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} + \lambda\pi_i + \pi_{i+1}\mu_{i+1}v'\left(\frac{y_i}{\theta_{i+1}}\right)\frac{1}{\theta_{i+1}} = 0 \text{ fo}$$

$$c_i \quad : \quad W_i\pi_i u'(c_i) + \pi_i\mu_i u'(c_i) - \lambda\pi_i = 0 \text{ for } i = I$$

$$y_i \quad : \quad -W_i\pi_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} - \pi_i\mu_i v'\left(\frac{y_i}{\theta_i}\right)\frac{1}{\theta_i} + \lambda\pi_i = 0 \text{ for } i = I$$

- $2I + I - 1 + 1$ unknowns: $\{c_i, y_i\}_{i=1}^{I}$, $\{\mu_i\}_{i=2}^{I}$, λ
- $2I$ FOCs + I constraints: $I - 1$ IC constraints and the resource constraint

Decentralization

- Problem will have a solution. How can we decentralize it?
- FOC and budget constraints for households

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

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

- Note that marginal and average tax rates are only exactly pinned down at grid points.
- Note that for $i = I$ the $i+1$ terms are absent, so can combine the two FOCs to give

$$\left(W_I\pi_I + \pi_I\mu_I\right)\theta_I u'(c_I) = \left(W_I\pi_I + \pi_I\mu_I\right)v'\left(\frac{y_I}{\theta_I}\right)$$

- $\Rightarrow T'(y_I) = 0$
- A classic result in the literature

Numerical Solution

1. Guess λ
2. Guess c_1
3. Solve for μ_2 from FOC for c_1
4. Solve for y_1 from FOC for y_1
5. 3 equations (2 FOCs and IC_2) to solve for c_2, y_2 and μ_3
6. Iterate upwards through the grid
7. At $I - 1$ we solve for μ_I
8. Then we have 2 FOCs at I to solve for c_I and y_I
9. Check IC_I and adjust c_1 if not satisfied
10. Finally check resource constraint and adjust λ

Practical Optimal Income Taxation (2021)

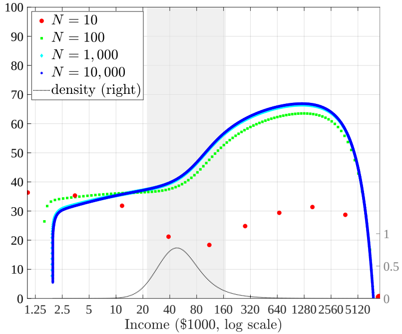
- In this paper we show that a very fine grid on productivity is required to deliver practical policy advice $\Rightarrow$ Policy prescriptions based on analyses with a coarse grid are of little practical value
- If using a fine grid is not feasible, a parametric Ramsey-style approach to optimal policy is preferable to the Mirrleesian approach

Calibration

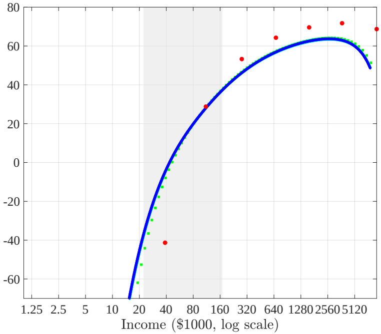
- Preferences: $\sigma = 2$
- G is equal to 18.8 percent of model GDP
- Wage distribution estimated using the SCF (will return to this)
- Discretization: N grid points, evenly spaced in logs such that $\{\log(\theta_1), \cdots, \log(\theta_N)\}$, with prob $\{\pi_1, \cdots, \pi_N\}$
- Coarseness of the grid: $\kappa = \frac{\theta_{i+1}}{\theta_i} = \left(\frac{\theta_N}{\theta_1}\right)^{\frac{1}{N-1}}$

Optimal Tax Policy and Grid Points

(a) Marginal Tax Rates



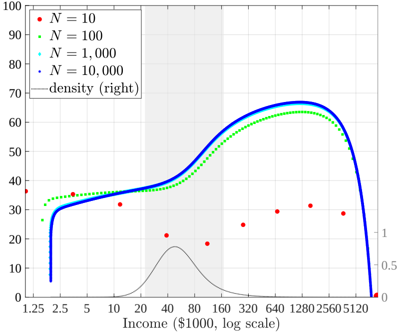
(b) Average Tax Rates



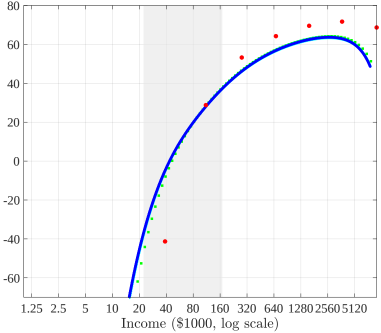
- With $N \geq 1,000$, tax schedules indistinguishable
  $\Rightarrow$ Accurate representation of optimal taxation in an economy with continuous productivity distribution

Optimal Tax Policy and Grid Points

(a) Marginal Tax Rates



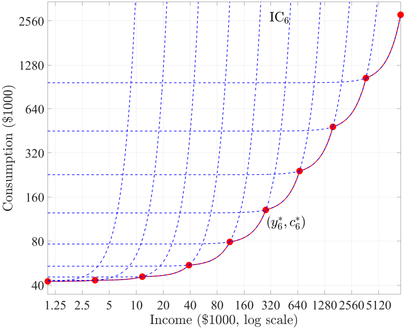
(b) Average Tax Rates



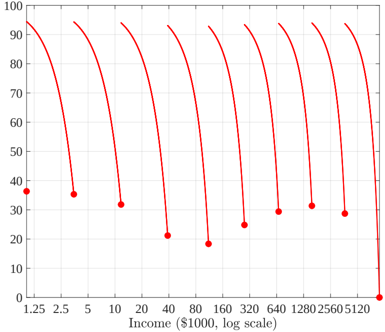
- With $N = 10$, MTRs lower, but ATR rises faster?
- Need to fill in taxes in between grid points

True Optimal Tax Policy when $N = 10$

(a) Income vs Consumption



(b) Marginal Tax Rates



Linearly interpolating between grid points would be wrong!

Welfare Gains

	Welfare Gains (%,CEV)		
	(1)	**(2)**	**(3)**
# of grid points N	Status Quo	First Best	Mirrlees
$10,000$	—	44.72	2.07
$1,000$	0.00	44.73	2.28
100	−0.01	44.81	4.40
50	−0.01	44.89	6.66
10	−0.21	46.07	20.13

Computing Optimal Taxes: Tax Formula Approach

- Take FOCs from Mirrlees problem and rearrange to get famous Diamond Saez tax formula equation

$$\frac{T_i^{*\prime}}{1-T_i^{*\prime}} = \underbrace{\frac{1-\kappa^{-(1+\sigma)}}{\kappa-1}}_{A} \times \underbrace{\frac{\kappa-1}{\pi_i}\sum_{s=i+1}^{N}\pi_s\left(1-\frac{\mathbb{E}\left[c\right]}{c_s}\right)\frac{c_s}{c_i}}_{B}$$

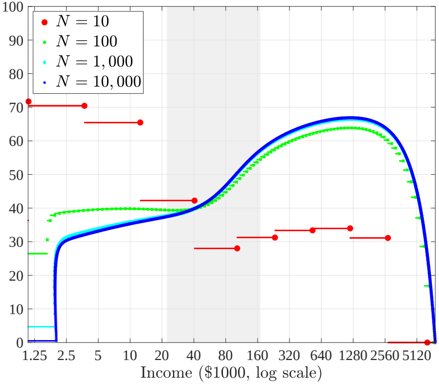
- In continuous time we get following more familiar version

$$\frac{T'\left(y(\theta)\right)}{1-T'\left(y(\theta)\right)} = \underbrace{(1+\sigma)}_{A} \times \underbrace{\frac{1}{\theta f(\theta)}\int_{\theta}^{\infty}\left(1-\frac{\mathbb{E}[c]}{c(s)}\right)\frac{c(s)}{c(\theta)}dF(s)}_{B}$$

- Use the DS formula to jointly solve for optimal taxes and allocations $\Rightarrow$ a fixed point problem

Computing Optimal Taxes: Tax Formula Approach

- Typical assumption: T piecewise linear (Mankiw et al)
- Guess MTR for each productivity type, solve for CE
- Use DS formula to update guesses



Practical Alternative: Flexible Ramsey Taxation

- Tax formula approach works fine if the grid is very fine
- With a coarse grid it fails – allocation the approach converges to does not solve the Mirrlees problem
- Method is doomed to fail because the optimal schedule is not piecewise linear

- So use a very fine grid, if possible
- But what if a fine enough grid is infeasible?
- We want an alternative approach that
  1. delivers a prescription close to true Mirrleesian optimum
  2. delivers similar prescription irrespective of coarseness
  3. is fast and easy to compute

Flexible Ramsey taxation

- Polylogarithmic marginal tax

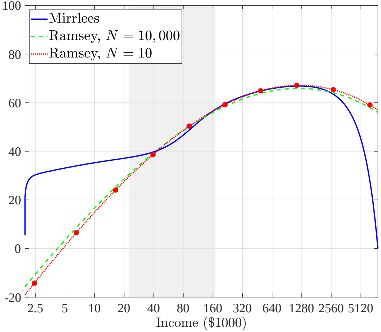
$$T'(y) = \sum_{i=1}^{M} \tau_i \left(\log y\right)^{i-1},$$

  with lump-sum taxes or transfers (ϕ_0)

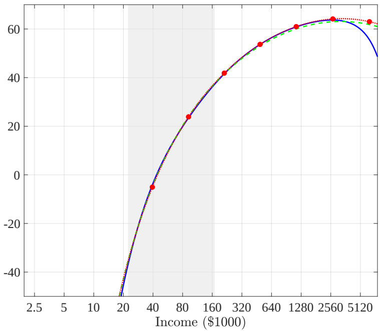
- Consider $M = 4$ so that we only search for $\{\phi_0, \tau_1, \tau_2, \tau_3, \tau_4\}$

Ramsey versus Mirrlees

(a) Marginal Tax Rates



(b) Average Tax Rates



Welfare gains: Ramsey optimum 2.0% and 1.9% for $N = 10,000$ and $N = 10$, respectively, compared to 2.1% under the true Mirrleesian optimum

Conclusions

- To characterize the optimum tax and transfer schedule, use a very fine grid.
- Not feasible? Go with Ramsey!