Optimal Income Taxation: Mirrlees Meets Ramsey

Jonathan Heathcote
FRB Minneapolis and CEPR

Hitoshi Tsujiyama
Goethe University Frankfurt

CIGS Conference, May 30-31, 2016

How should we tax income?

- What structure of income taxation offers best trade-off between benefits of public insurance and costs of distortionary taxes?
- Proposals for a flat tax system with universal transfers
  - Friedman (1962)
  - Mirrlees (1971)
- Others have argued for U-shaped marginal tax schedule
  - Saez (2001)

This Paper

We compare 3 tax and transfer systems:

1. Affine tax system: $T(y) = \tau_0 + \tau_1 y$
    - constant marginal rates with lump-sum transfers
2. HSV tax system: $T(y) = y - \lambda y^{1-\tau}$
    - function introduced by Feldstein (1969), Persson (1983), and Benabou (2000)
    - increasing marginal rates without transfers
    - τ indexes progressivity: $1 - \tau = \frac{1-T'(y)}{1-T(y)/y}$
3. Optimal tax system
    - fully non-linear

Main Findings

- Marginal tax rates should be increasing in income, NOT flat or U-shaped
- Best tax and transfer system in the HSV class typically better than the best affine tax system
  - More valuable to have marginal tax rates increase with income than to have lump-sum transfers
- Welfare gains from tax reform sensitive to planner's taste for redistribution
  - May be tiny

Mirrlees Approach to Tax Design: Mirrlees (1971), Diamond (1988), Saez (2001)

- Agents differ wrt unobservable log productivity α
- Planner only observes earnings $x = \exp(\alpha) \times h$
- Think of planner choosing (c, x) for each α type
- Include incentive constraints, s.t. each type prefers the earnings level intended for their type
- Allocations are constrained efficient
- Trace out tax decentralization $T(x(\alpha)) = x(\alpha) - c(\alpha)$

Novel Elements of Our Analysis

1. We explore a range of Social Welfare Functions
   - Utilitarian SWF as a benchmark
     $\Rightarrow$ Strong desire for redistribution
   - Alternative SWF that rationalizes amount of redistribution embedded in observed tax system
2. Our model has a distinct role for private insurance
   - Standard decentralization of efficient allocations delivers all insurance through tax system $\Rightarrow$ Very progressive taxes

Environment 1

- Standard static Mirrlees plus partial private insurance (quantitatively important)
- Heterogeneous individual labor productivity with two stochastic components

$$\log w = \alpha + \varepsilon$$

- ε is privately-insurable, α is not
    - Agents belong to large families
    - α common across all members of a family $\Rightarrow$ cannot be pooled within family
    - ε purely idiosyncratic & orthogonal to α $\Rightarrow$ can be pooled within family
- Planner sees neither component of productivity

Environment 2

- Common preferences

$$u(c,h) = \log(c) - \frac{h^{1+\sigma}}{1+\sigma}$$

- Production linear in aggregate effective hours

$$\int\int \exp(\alpha+\varepsilon)h(\alpha,\varepsilon)dF_\alpha dF_\varepsilon = \int\int c(\alpha,\varepsilon)dF_\alpha dF_\varepsilon + G$$

Planner's Problems

- Seeks to maximize SWF denoted $W(\alpha)$
- Only sees total family income $y(\alpha) = \int \exp(\alpha + \varepsilon) h(\alpha, \varepsilon) dF_\varepsilon$
- First Stage
  - Planner offers menu of contracts $\{c(\widetilde{\alpha}), y(\widetilde{\alpha})\}$
  - Family heads draw idiosyncratic α and report $\widetilde{\alpha}$
- Second Stage
  - Family members draw idiosyncratic ε
  - Family head tells each member how much to work
  - Total earnings must deliver $y(\widetilde{\alpha})$ to the planner
  - Must divide consumption $c(\widetilde{\alpha})$ between family members

Nature of the Solution

- Planner cannot condition individual allocations on ε, given free within-family transfers
  - equally cheap for any family member to deliver income to the planner, and equally valuable to receive consumption
- Thus, planner cannot take over private insurance

  $\Rightarrow$ Distinct roles for public and private insurance
- Note: Extent of private risk-sharing is exogenous with respect the tax system

Planner's Problem: Second Best

$$\max_{c(\alpha),y(\alpha)} \quad \int W(\alpha)U(\alpha,\alpha)dF_\alpha$$

$$\text{s.t.} \quad \int y(\alpha)dF_\alpha \geq \int c(\alpha)dF_\alpha + G$$

$$U(\alpha,\alpha) \geq U(\alpha,\widetilde{\alpha}) \qquad \forall\alpha, \forall\widetilde{\alpha}$$

where $U(\alpha,\widetilde{\alpha}) \equiv$

$$\begin{cases} \max_{\{c(\alpha,\tilde{\alpha},\varepsilon),h(\alpha,\tilde{\alpha},\varepsilon)\}} \int \left\{ \log(c(\alpha,\tilde{\alpha},\varepsilon)) - \frac{h(\alpha,\tilde{\alpha},\varepsilon)^{1+\sigma}}{1+\sigma} \right\} dF_\varepsilon \\ \text{s.t.} \quad \int c(\alpha,\tilde{\alpha},\varepsilon)dF_\varepsilon = c(\tilde{\alpha}) \\ \int \exp(\alpha+\varepsilon)h(\alpha,\tilde{\alpha},\varepsilon)dF_\varepsilon = y(\tilde{\alpha}) \end{cases}$$

$$U(\alpha,\widetilde{\alpha}) = \log(c(\widetilde{\alpha})) - \frac{\Omega}{1+\sigma}\left(\frac{y(\widetilde{\alpha})}{\exp(\alpha)}\right)^{1+\sigma}$$

$$\text{where} \quad \Omega = \left(\int \exp(\varepsilon)^{\frac{1+\sigma}{\sigma}} dF_\varepsilon(\varepsilon)\right)^{-\sigma}$$

Planner's Problem: Ramsey

$$\max_{\tau} \quad \int W(\alpha)\left\{\int u(c(\alpha,\varepsilon),h(\alpha,\varepsilon))dF_\varepsilon\right\}dF_\alpha$$

$$\text{s.t.} \quad \int\int c(\alpha,\varepsilon)dF_\alpha dF_\varepsilon + G = \int\int \exp(\alpha+\varepsilon)h(\alpha,\varepsilon)dF_\alpha dF_\varepsilon$$

where $c(\alpha,\varepsilon)$ and $h(\alpha,\varepsilon)$ are the solutions to

$$\begin{cases} \max_{\{c(\alpha,\varepsilon),h(\alpha,\varepsilon)\}} & \int\left\{\log c(\alpha,\varepsilon) - \frac{h(\alpha,\varepsilon)^{1+\sigma}}{1+\sigma}\right\}dF_\varepsilon \\ \text{s.t.} & \int c(\alpha,\varepsilon)dF_\varepsilon = y(\alpha) - T\left(y(\alpha);\tau\right) \\ & y(\alpha) = \int \exp(\alpha+\varepsilon)h(\alpha,\varepsilon)dF_\varepsilon \end{cases}$$

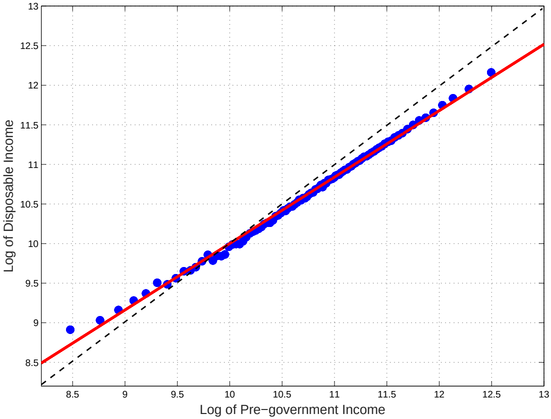
Social Preferences

- Assume SWF takes the form $W(\alpha;\theta) = \exp(-\theta\alpha)$
  - θ controls taste for redistribution
  - $W(\alpha;\theta)$ function could be micro-founded as a probabilistic voting model
- Nests standard SWFs used in the literature:
  - $\theta = 0$: Utilitarian [our benchmark]
  - $\theta = -1$: Laissez-Faire Planner
  - $\theta \to \infty$: Rawlsian

Empirically Motivated SWF

- Progressivity built into current tax system informative about politico-economic demand for redistribution
- Assume planner (political system) choosing tax system in HSV class: $T(y) = y - \lambda y^{1-\tau}$
- Assume planner has SWF in class $W(\alpha;\theta) = \exp(-\theta\alpha)$
- What value for θ gives observed τ as solution to Ramsey problem?
    - Let $\tau^*(\theta)$ denote welfare-maximizing choice for τ given θ
    - Empirically Motivated SWF $W(\alpha;\theta^*)$ s.t. $\tau^*(\theta^*) = \tau^{US}$
    - related to inverse optimum problem
- Ramsey planner with $\theta = \theta^*$ choosing a tax and transfer scheme in the HSV class would choose exactly τ^{US}

Baseline HSV Tax System: $T(y; \lambda, \tau) = y - \lambda y^{1-\tau}$

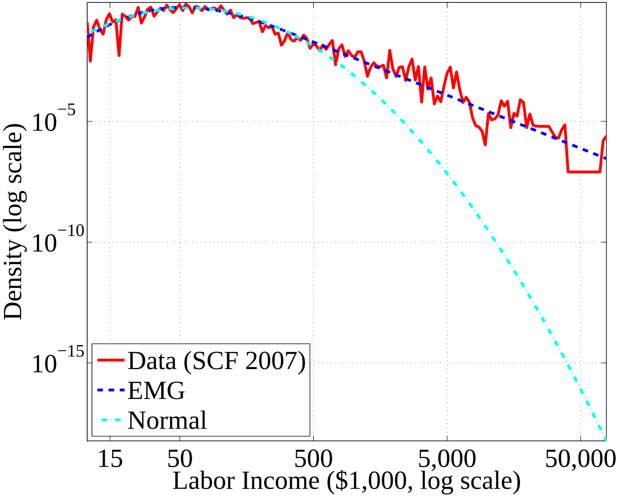


- Estimated on PSID data for 2000-2006
- Households with head / spouse hours $\geq$ 260 per year
- Estimated value for $\tau = 0.161$, $R^2 = 0.96$

Calibration: Wage Distribution

- Heavy Pareto-like right tail of labor earnings distribution (Saez, 2001)
- Assume Pareto tail reflects uninsurable wage dispersion
- F_α : Exponentially Modified Gaussian $EMG(\mu_\alpha, \sigma_\alpha^2, \lambda_\alpha)$
- F_ε : Normal $N(\frac{-\sigma_\varepsilon^2}{2}, \sigma_\varepsilon^2)$
- $\log(w) = \alpha + \varepsilon$ is itself EMG $\Rightarrow w$ is Pareto log-normal
- $\log(wh)$ is also EMG, given our utility function, private insurance model, and HSV tax system
- Normal variance coefficient in the EMG distribution for log earnings: $\sigma_y^2 = \left(\frac{1+\sigma}{\sigma+\tau}\right)^2 \sigma_\varepsilon^2 + \sigma_\alpha^2$.

Distribution for Labor Income



Use micro data from the 2007 SCF to estimate α by maximum likelihood $\Rightarrow$ $\lambda_\alpha = 2.2$ and $\sigma_y^2 = 0.4117$

Calibration

- Frisch elasticity $= 0.5 \Rightarrow \sigma = 2$

- Progressivity parameter $\tau = 0.161$ (HSV 2014)
- Govt spending G s.t. $G/Y = 0.188$ (US, 2005)

- Variance of normal component of SCF earnings + external evidence on importance of insurable shocks
  $\Rightarrow \sigma_\varepsilon^2 = \sigma_\alpha^2 = 0.1407$

  - Variance of insurable shocks consistent with HSV 2014

  - Total variance of log wages (0.488) and variance of log consumption (0.246) consistent with empirical counter parts

Bottom of Wage Distribution

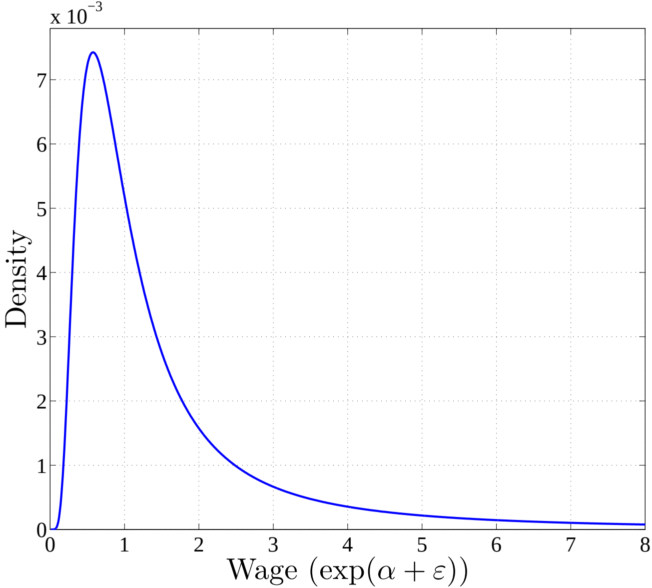
- Difficult to measure distribution of offered wages at the bottom, given selection into participation
- Low and Pistaferri (2015) estimate distribution of latent offered wages within a structural model in which workers face disability risk and choose participation

Percentile Ratios	Model	LP
P5/P1	1.48	1.48
P10/P5	1.24	1.20
P25/P10	1.44	1.40

Numerical Implementation

- Maintain continuous distribution for ε
- Assume a discrete distribution for α
- Baseline: 10,000 evenly-spaced grid points
- α_{min}: \$2 per hour (5% of the average $=$ \$41.56)
- α_{max}: \$3,075 per hour (\$6.17m assuming 2,000 hours $=$ 99.99th percentile of SCF earnings distn.)
- Set μ_α and σ_α^2 to match $E\left[e^\alpha\right] = 1$ and target for $var(\alpha)$ given $\lambda_\alpha = 2.2$

Wage Distribution



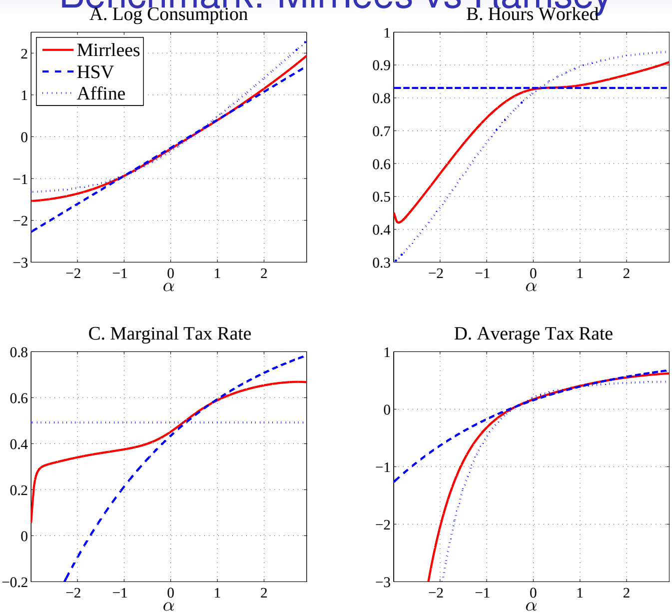
Quantitative Analysis

- U.S. tax system approximated by HSV with $\tau = 0.161$
- Focus on three optimal systems:
  1. HSV tax function: $T(y) = y - \lambda y^{1-\tau}$
  2. Affine tax function: $T(y) = \tau_0 + \tau_1 y$
  3. Mirrless tax function (second best allocation)

Quantitative Analysis: Benchmark

Tax System	Tax Parameters		Outcomes			
			welfare	Y	$T'(y)$	TR/Y
HSV^{US}	$\lambda : 0.839$	$\tau : 0.161$	–	–	0.319	0.018
HSV	$\lambda : 0.817$	$\tau : 0.330$	2.08	-7.22	0.466	0.063
Affine	$\tau_0 : -0.259$	$\tau_1 : 0.492$	1.77	-8.00	0.492	0.279
Mirrlees			2.48	-7.99	0.491	0.213

Benchmark: Mirrlees vs Ramsey



Quantitative Analysis: Benchmark

- Optimal HSV better than optimal affine

  $\Rightarrow$ Increasing marginal rates more important than lump-sum transfers

- Moving to fully optimal system generates substantial gains (2.5%)

- The optimal marginal tax rate is around 50%

Quantitative Analysis: Sensitivity

What drives the results?

1. Eliminate insurable shocks: $\tilde{v}_\alpha = v_\alpha + v_\varepsilon$ and $\tilde{v}_\varepsilon = 0$
2. Utilitarian SWF $\theta = 0$

   $\Rightarrow$ Various SWFs including Empirically motivated SWF
3. Increase desire to raise revenue
4. Wage distribution has thin Log-Normal right tail: $\alpha \sim N$

Sensitivity: No Insurable Shocks

Tax System	Tax Parameters		Outcomes			
			welfare	Y	$T'(y)$	TR/Y
HSV^{US}	$\lambda : 0.842$	$\tau : 0.161$	–	–	0.319	0.019
HSV	$\lambda : 0.804$	$\tau : 0.383$	4.17	-9.72	0.511	0.084
Affine	$\tau_0 : -0.283$	$\tau_1 : 0.545$	5.34	-10.45	0.545	0.326
Mirrlees			5.74	-10.64	0.550	0.284

- No insurable shocks $\Rightarrow$ larger role for public redistribution
- Want higher tax rates and larger transfers
- Optimal HSV worse than optimal affine

  $\Rightarrow$ Distinguishing insurable shocks from uninsurable shocks is important

Social Welfare

- Consider alternative SWFs:
  - $\theta = -1$: Laissez-Faire Planner
  - $\theta \to \infty$: Rawlsian
- Empirically motivated SWF: $W(\alpha; \theta^*)$ s.t. $\tau^*(\theta^*) = \tau^{US}$
  - Closed form expression for θ^*!

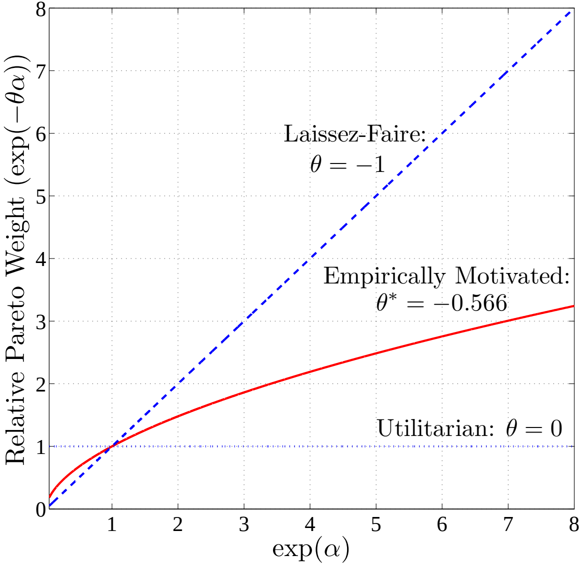
$$\sigma_\alpha^2 \theta^* - \frac{1}{\lambda_\alpha + \theta^*} = -\frac{1}{\lambda_\alpha - 1 + \tau} - \sigma_\alpha^2 (1 - \tau) + \frac{1}{1+\sigma} \left\{ \frac{1}{(1-g)(1-\tau)} - 1 \right\}$$

- Simple in Normal case ($\lambda_\alpha \to \infty$)

$$\theta^* = -(1 - \tau) + \frac{1}{\sigma_\alpha^2} \frac{1}{1 + \sigma} \left\{ \frac{1}{(1 - g)(1 - \tau)} - 1 \right\}$$

  - θ^* increasing in τ and g
  - θ^* declining in σ and σ_α^2
  - θ^* increasing in λ_α (holding fixed $var(\alpha) = \sigma_\alpha^2 + \frac{1}{\lambda_\alpha^2}$)

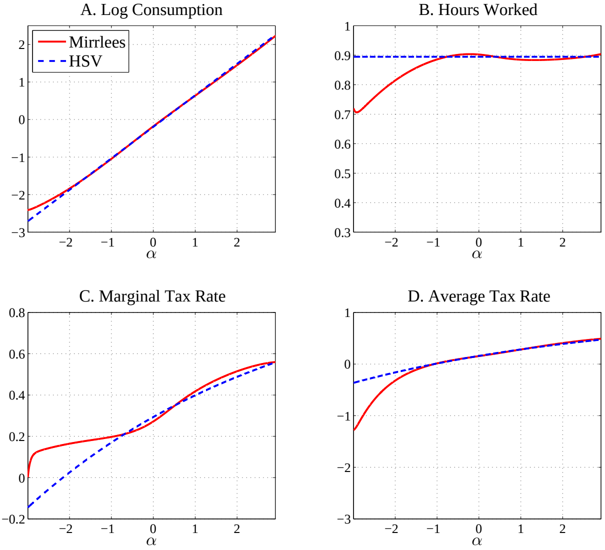
Social Welfare Functions



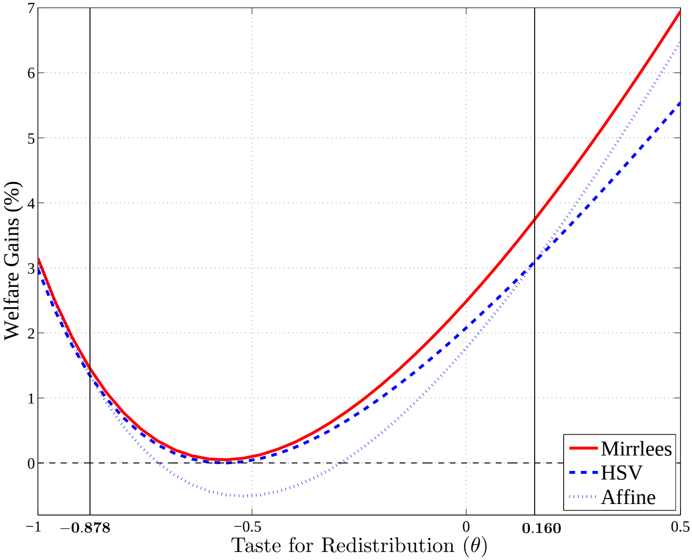
Sensitivity: Alternative SWFs

SWF		Mirrlees Allocations			Welfare Change		
	θ	$T'(y)$	TR/Y	ΔY	Mirrlees	Affine	HSV
Laissez-Faire	-1	0.083	-0.082	9.72	3.15	3.14	2.98
Emp. Motivated	-0.57	0.314	0.051	0.16	0.05	-0.48	–
Utilitarian	0	0.491	0.213	-7.99	2.48	1.77	2.08
Rawlsian	∞	0.711	0.538	-22.55	708.28	649.14	354.90

Empirically-Motivated SWF



HSV vs Affine with Various SWFs



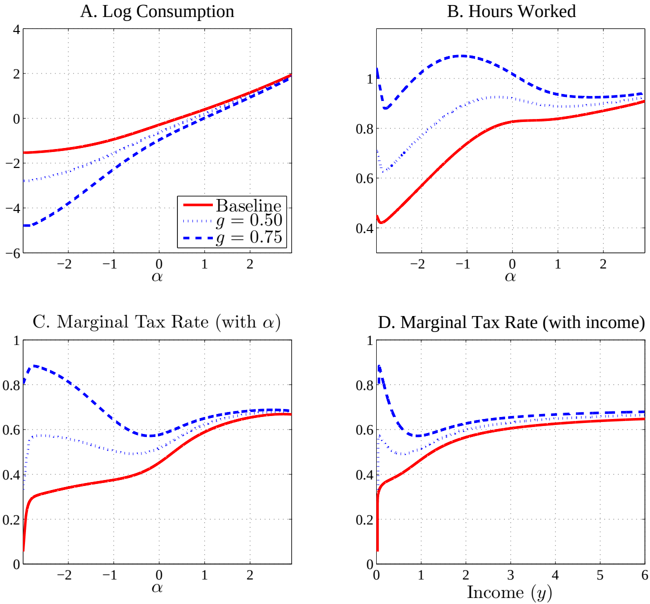
SWF Sensitivity: Summary

- Optimal tax system very sensitive to assumed SWF
- Welfare gains moving from the current tax system to the optimal one can be tiny
- Affine system works well when preference for redistribution is either very strong or very weak:
  - In the first case, want large lump-sum transfers
  - In the second, want lump-sum taxes
- For intermediate tastes for redistribution ($\theta \in [-0.88, 0.16]$), HSV is better than affine

Sensitivity: Need to Raise Revenue

- Saez (2001) found a U-shaped marginal schedule to be optimal
- His intuition: Want to make sure welfare is targeted only to the very poor
- We don't find this. Why?
- Key is degree of revenue requirement: to finance
  - exogenous public expenditure G
  - endogenous universal lump-sum transfers Tr

U-shaped Tax Rates with High G



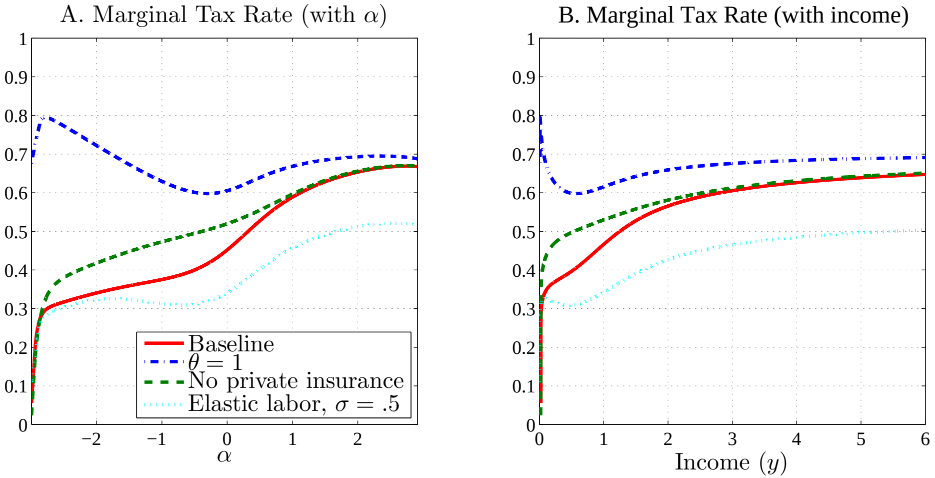
Intuition: U-shaped Tax Rates with High G

- Tax rates at the top relatively insensitive to the level of G
  - Already close to the top of the Laffer curve
  - Asymptotic rates indicated by Saez (2001): $\frac{1+\sigma}{\sigma+\lambda_\alpha} \approx 71\%$
- Tax rates at low income levels increase in G
  - Little room at the top $\Rightarrow$ instead raise marginal rates at low income levels
- U-shaped rather than monotonically declining
  - Dip in the middle to keep labor supply distortions low where the heaviest population mass is located

Alternative Ways to Increase Fiscal Pressure

- Increase optimal lump-sum transfers by
  - Increasing the planner's taste for redistribution $\theta = 1$
  - Shutting off private insurance
- Reduce the government's ability to satisfy revenue demands by
  - Increasing the labor supply elasticity $\sigma = 0.5$

Alternative Ways to Increase Fiscal Pressure



Why does Saez (2001) find U-shaped rates?

- Various assumptions that imply high fiscal pressure:
  - Higher value for government purchases (25% of GDP)
  - Rule out private insurance
  - Use utility functions that limit the government's ability to extract revenue from the rich
- U-shaped profile for marginal rates is not a general feature of an optimal tax system

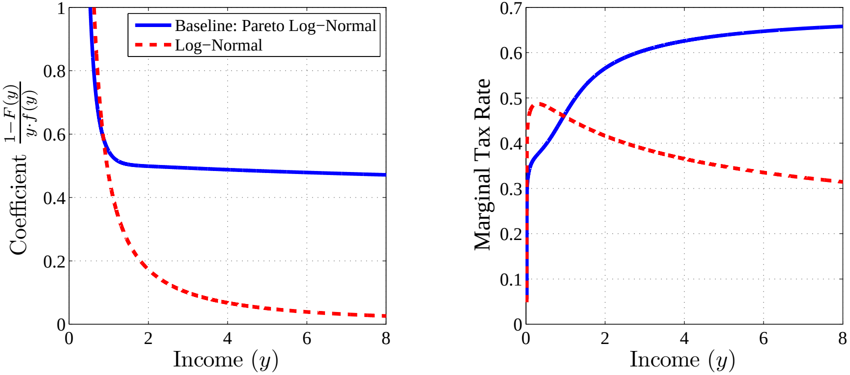
Sensitivity: Log-Normal Wage

Tax System	Tax Parameters		Outcomes			
			welfare	Y	$T'(y)$	TR/Y
HSV^{US}	$\lambda : 0.828$	$\tau : 0.161$	–	–	0.319	0.017
HSV	$\lambda : 0.813$	$\tau : 0.285$	0.88	-5.20	0.427	0.048
Affine	$\tau_0 : -0.230$	$\tau_1 : 0.451$	2.19	-6.01	0.451	0.242
Mirrlees			2.28	-5.74	0.443	0.254

- Log-normal distribution $\Rightarrow$ thin right tail
- Optimal HSV worse than optimal affine
- Optimal affine nearly efficient

Why Distribution Shape Matters

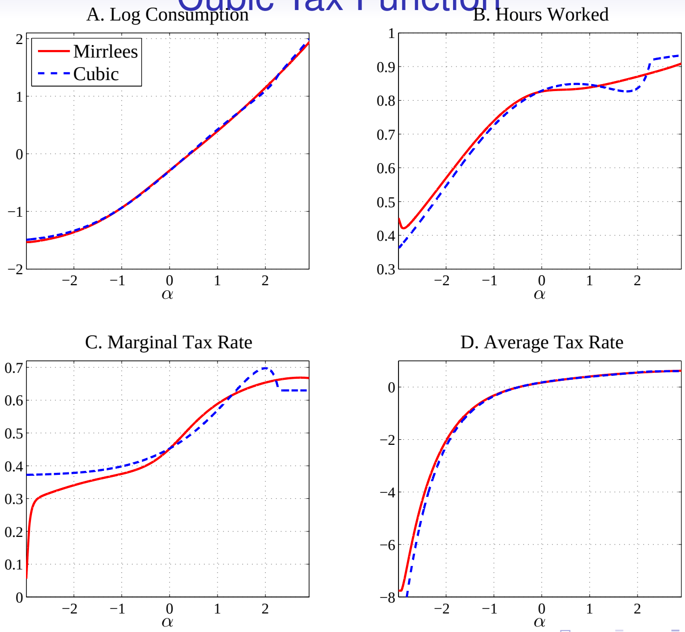
- Want high top marginal rates when (i) few agents face those marginal rates, but (ii) can capture lots of revenue from higher-income households



Extension: Polynomial Tax Functions

Tax System	Tax Parameters				Outcomes			
					welfare	Y	$T'(y)$	TR/Y
HSV^{US}	λ 0.839	τ 0.161			–	–	0.319	0.018
Affine	τ_0 −0.259	τ_1 0.492			1.77	−8.00	0.492	0.279
Cubic	τ_0 −0.212	τ_1 0.370	τ_2 0.049	τ_3 −0.002	2.40	−8.01	0.491	0.228
Mirrlees					2.48	−7.99	0.491	0.213

Cubic Tax Function



Extension: Type-Contingent Taxes

- Productivity partially reflects observable characteristics (e.g. education, age, gender)
- Some fraction of uninsurable shocks are observable: $\alpha \to \alpha + \kappa$
- Heathcote, Perri & Violante (2010) estimate variance of cross-sectional wage dispersion attributable to observables, $v_\kappa = 0.108$
- Planner should condition taxes on observables: $T(y; \kappa)$
- Consider two-point distribution for κ (college vs high school)

Extension: Type-Contingent Taxes

- Significant welfare gains relative to non-contingent tax
- Conditioning on observables ⇒ marginal tax rates of 42%

System		Outcomes			
		wel.	Y	$T'(y)$	TR/Y
HSV^{US}	$\lambda : 0.834, \tau : 0.161$	–	–	0.319	0.015 0.020
HSV	$\lambda^L : 1.069, \tau^L : 0.480$ $\lambda^H : 0.595, \tau^H : 0.073$	6.21	−2.80	0.416	0.147 −0.019
Affine	$\tau_0^L : -0.403, \tau_1^L : 0.345$ $\tau_0^H : -0.032, \tau_1^H : 0.452$	6.15	−2.53	0.421	0.420 0.008
Mirrlees		6.54	−2.53	0.418	0.368 0.007

Conclusions

- Optimal marginal tax schedule increasing in income, and neither flat nor U-shaped
- Welfare gains moving from the current tax system to the optimal one hinge on the choice of SWF, may be tiny
- Ramsey and Mirrlees tax schemes not far apart: can approximately decentralize Mirrlees with a simple tax scheme