Optimal Income Taxation: Mirrlees Meets Ramsey

Jonathan Heathcote
FRB Minneapolis and CEPR

Hitoshi Tsujiyama
Goethe University Frankfurt

ESSIM May 28 2014

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System

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)

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: $T(y) = y - \lambda y^{1-\tau}$
   - functional form introduced by Feldstein (1969), Persson (1983), and Benabou (2000)
   - increasing marginal rates without transfers
   - τ indexes progressivity of the system
3. Optimal tax system
   - fully non-linear

Main Findings

- Best tax and transfer system in the HSV class better than the best affine tax system
- Welfare gains moving from the current tax system to the optimal one are tiny

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

- Planner only observes earnings = productivity $\times$ effort
- Think of planner choosing earnings x and cons. c for each unobservable productivity type α
- 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. Our model has a distinct role for private insurance
   - Standard decentralization of efficient allocations delivers all insurance through tax system $\Rightarrow$ Very progressive taxes
2. We use a SWF that rationalizes amount of redistribution embedded in observed tax system
   - Analyzes typically assume utilitarian social welfare function $\Rightarrow$ Strong desire for redistribution

Environment 1

- Static environment
- Heterogeneous individual labor productivity w
- Log productivity is sum of two independent stochastic components

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

- α no private insurance
- ε private insurance
- planner sees neither component of productivity
    - (later introduce a third productivity component κ that the planner can observe)

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 $W(\alpha, \varepsilon)$
- Only observes total income y = earnings plus private insurance income
- First Stage
  - Planner offers menu of contracts $\{c(\widetilde{\alpha}, \widetilde{\varepsilon}), y(\widetilde{\alpha}, \widetilde{\varepsilon})\}$
  - Agents draw idiosyncratic α and report $\widetilde{\alpha}$
- Second Stage
  - Agents buy private insurance against insurable shock ε
  - Draw ε, receive insurance payments and report $\widetilde{\varepsilon}$
  - Work sufficient hours to deliver $y(\widetilde{\alpha}, \widetilde{\varepsilon})$
  - Receive consumption $c(\widetilde{\alpha}, \widetilde{\varepsilon})$

First result: Cannot condition on $\widetilde{\varepsilon}$

- Offered contracts take the form $\{c(\widetilde{\alpha}), y(\widetilde{\alpha})\}$
- Private insurance markets undercut planner's ability to condition allocations on ε
- Planner cannot take over private insurance $\Rightarrow$ Distinct roles for public and private insurance

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

$$\left\{ \begin{array}{ll} \max_{h(\varepsilon),B(\varepsilon)} & \int \left\{ \log c(\widetilde{\alpha}) - \frac{h(\varepsilon;\alpha,\widetilde{\alpha})^{1+\sigma}}{1+\sigma} \right\} dF_\varepsilon \\ \text{s.t.} & \int Q(\varepsilon)B(\varepsilon;\alpha,\widetilde{\alpha})d\varepsilon = 0 \\ & \exp(\alpha+\varepsilon)h(\varepsilon;\alpha,\widetilde{\alpha}) + B(\varepsilon;\alpha,\widetilde{\alpha}) = y(\widetilde{\alpha}) \quad \forall\varepsilon \end{array} \right.$$

price of insurance $\int_E Q(\varepsilon)d\varepsilon = \int_E dF_\varepsilon$

Planner's Problem: Ramsey

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

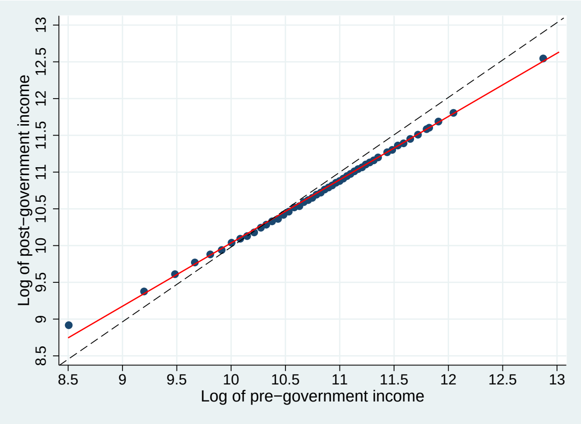
$$\text{s.t.} \quad \int\int y(\alpha,\varepsilon)dF_\alpha dF_\varepsilon \geq \int\int c(\alpha,\varepsilon)dF_\alpha dF_\varepsilon + G$$

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

$$\begin{cases} \max\limits_{c(\alpha,\varepsilon),h(\alpha,\varepsilon),B(\alpha,\varepsilon)} & \int \left\{\log(c(\alpha,\varepsilon)) - \frac{h(\alpha,\varepsilon)^{1+\sigma}}{1+\sigma}\right\} dF_\varepsilon \\ \text{s.t.} & \int Q(\varepsilon)B(\alpha,\varepsilon)d\varepsilon = 0 \\ & c(\alpha,\varepsilon) \leq y(\alpha,\varepsilon) - T(y(\alpha,\varepsilon);\tau) \qquad \forall\varepsilon \end{cases}$$

where $y(\alpha,\varepsilon) \equiv \exp(\alpha+\varepsilon)h(\alpha,\varepsilon) + B(\alpha,\varepsilon)$

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

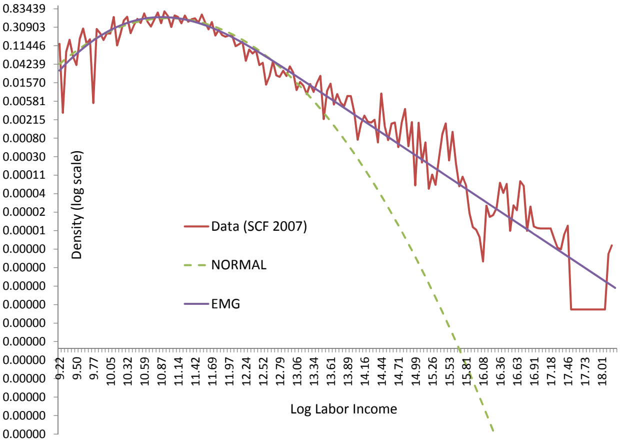


- Estimated on PSID data for 2000-2006
- Households with head / spouse hours $\geq$ 260 per year
- Estimated value for $\tau = 0.151$, $R^2 = 0.96$

Baseline 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 $EMG(\mu, \eta^2, a)$
- F_ε : Normal $N(\frac{-v_\varepsilon}{2}, v_\varepsilon)$
- $\log(w) = \alpha + \varepsilon$ is itself EMG $\Rightarrow w$ is Pareto-Lognormal
- $\log(wh)$ is also EMG, given our utility function, market structure, and HSV tax system

Distribution for Labor Income



Use micro data from the 2007 SCF to estimate α by maximum likelihood $\Rightarrow$ $a = 2.2$

Baseline Social Welfare Function

- Progressivity built into current tax system informative about society's taste for redistribution
- Assume SWF takes the form

$$W(\alpha) = \exp(-\theta\alpha)$$

- θ controls taste for redistribution (e.g. $\theta = 0$: utilitarian)
- Assume govt choosing a tax system in HSV class

$$T(y) = y - \lambda y^{1-\tau}$$

- What value for θ rationalizes observed choice for τ?
- Empirically-Motivated SWF: θ^{US} that solves $\tau^*(\theta^{US}) = \tau^{US}$

Social Welfare

- θ^{US} solves

$$-\eta^2\theta^{US}+\frac{1}{a+\theta^{US}} = \frac{1}{a-1+\tau^{US}}+\eta^2(1-\tau^{US})+\frac{1}{1+\sigma}\left\{1-\frac{1}{(1-g^{US})(1-\tau^{US})}\right\}$$

  - θ^{US} is increasing in τ and g
  - θ^{US} is decreasing in η^2 and σ

- Special case: If F_α is also Normal ($a \to \infty$),

$$\theta^{US} = -(1-\tau^{US}) + \frac{1}{\eta^2}\frac{1}{1+\sigma}\left\{\frac{1}{(1-g^{US})(1-\tau^{US})}-1\right\}$$

- Use θ^{US} as baseline for welfare comparisons $\Rightarrow$ focus on relative efficiency of alternative tax systems

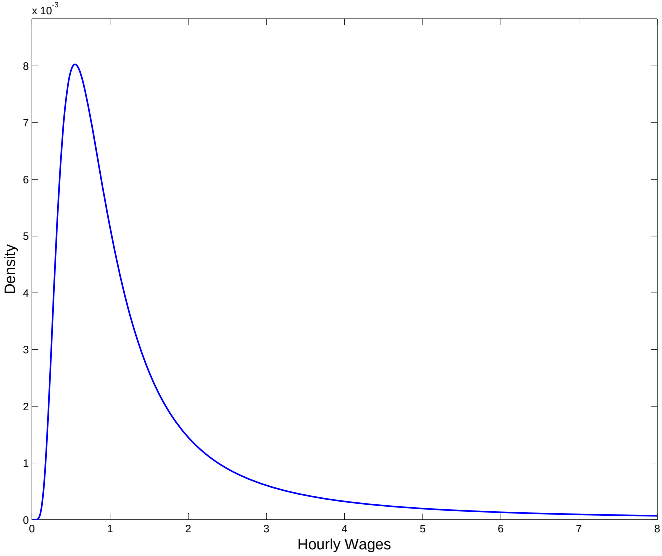
Calibration

- Frisch elasticity $= 0.5 \rightarrow \sigma = 2$
- Progressivity parameter $\tau = 0.151$ (HSV 2014)
- Govt spending G s.t. $G/Y = 0.188$ (US, 2005)
- $var(\varepsilon) = 0.193$: estimated variance of insurable shocks (HSV 2013)
- $var(\alpha) = 0.273$: total variance of wages is 0.466

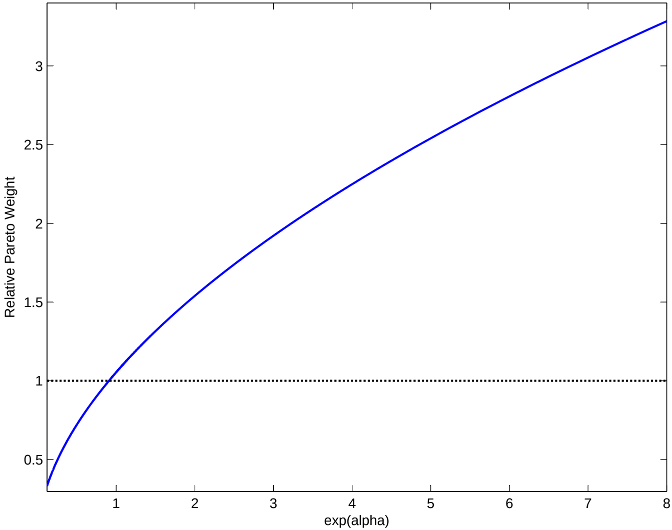
Numerical Implementation

- Maintain continuous distribution for ε
- Assume a discrete distribution for α
- Baseline: 10,000 evenly-spaced grid points
- $\alpha_{\min}$: \$5 per hour (12% of the average $=$ \$41.56)
- $\alpha_{\max}$: \$3,075 per hour (\$6.17m assuming 2,000 hours $=$ 99.99th percentile of SCF earnings distn.)
- Set μ and η^2 to match $E\left[e^{\alpha}\right] = 1$ and target for $var(\alpha)$ given $a = 2.2$

Wage Distribution



Social Welfare Function



Quantitative Analysis: Benchmark

Tax System	Tax Parameters				Outcomes			
					welfare	Y	mar. tax	TR/Y
HSV^{US}	λ 0.836	τ 0.151			–	–	0.311	0.018
Affine	τ_0 −0.116	τ_1 0.303			−0.58	0.41	0.303	0.089
Cubic	τ_0 −0.032	τ_1 0.126	τ_2 0.064	τ_3 −0.003	0.05	0.79	0.289	0.017
Mirrlees					0.11	0.82	0.287	0.003

Quantitative Analysis: Benchmark

- Moving to affine tax system is walfare reducing

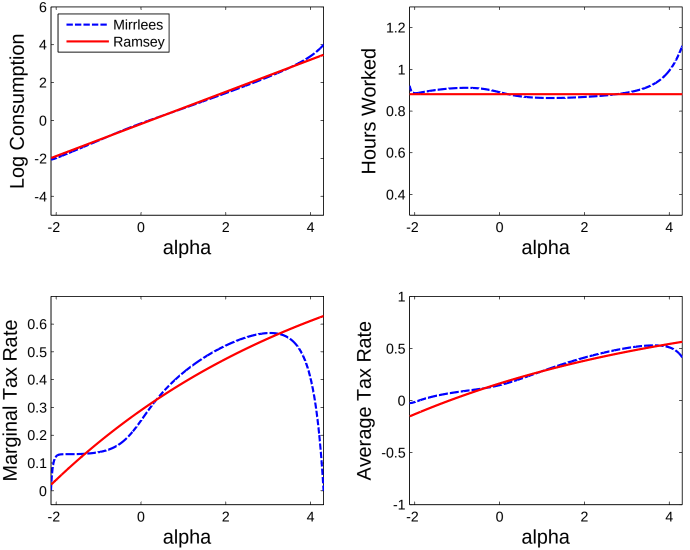
  $\Rightarrow$ Increasing marginal rates more important than lump-sum transfers

- Moving to fully optimal system generates only tiny gains (0.1%)

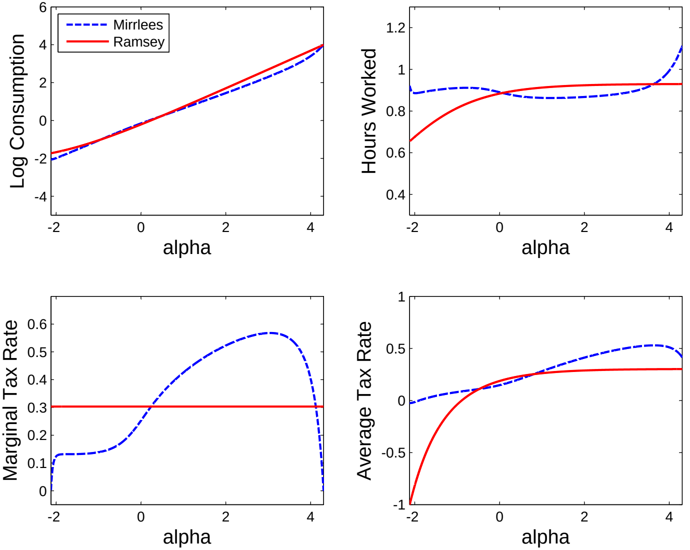
- The optimal marginal tax rate is around 30%

- Almost no need for transfers

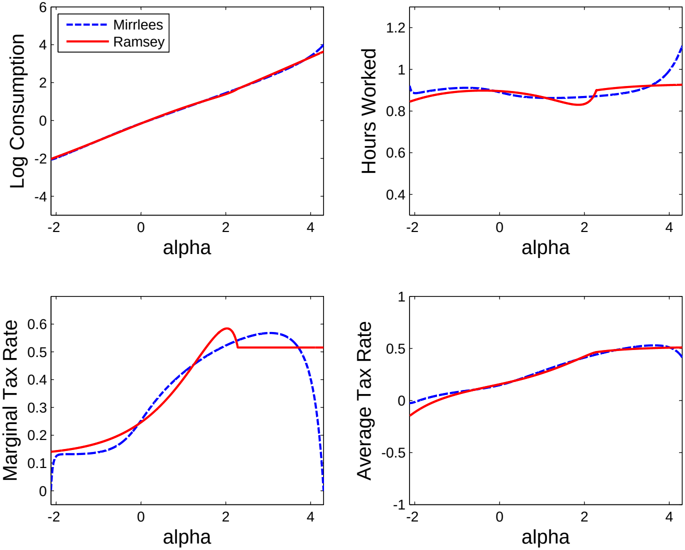
HSV Tax Function



Affine Tax Function



Cubic Tax Function



Quantitative Analysis: Sensitivity

What drives the results?

1. Empirically-motivated SWF $\rightarrow$ Utilitarian SWF: $\theta = 0$
2. Eliminate insurable shocks: $\tilde{v}_\alpha = v_\alpha + v_\varepsilon$ and $\tilde{v}_\varepsilon = 0$
3. Wage distribution has thin Log-Normal right tail: $\alpha \sim N$

Sensitivity: Utilitarian SWF

- Utilitarian SWF $\Rightarrow$ stronger taste for redistribution
- Want higher tax rates and larger transfers
- Optimal HSV still better than optimal affine

Tax System	Tax Parameters		Outcomes			
			welfare	Y	mar. tax	TR/Y
HSV^{US}	$\lambda : 0.836$	$\tau : 0.151$	–	–	0.311	0.018
HSV	$\lambda : 0.821$	$\tau : 0.295$	1.38	−6.02	0.436	0.068
Affine	$\tau_0 : -0.233$	$\tau_1 : 0.452$	0.45	−6.43	0.452	0.220
Mirrlees			1.53	−6.15	0.440	0.122

Sensitivity: No Insurable Shocks

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

Tax System	Tax Parameters		Outcomes			
			welfare	Y	mar. tax	TR/Y
HSV^{US}	$\lambda: 0.836$	$\tau: 0.151$	–	–	0.311	0.018
HSV	$\lambda: 0.839$	$\tau: 0.192$	0.12	-1.64	0.346	0.033
Affine	$\tau_0: -0.156$	$\tau_1: 0.360$	-0.21	-1.97	0.360	0.139
Mirrlees			0.23	-2.11	0.361	0.081

- Utilitarian SWF + No insurable shocks
  $\Rightarrow$ Lump-sum transfers more important
  $\Rightarrow$ Optimal HSV worse than optimal affine

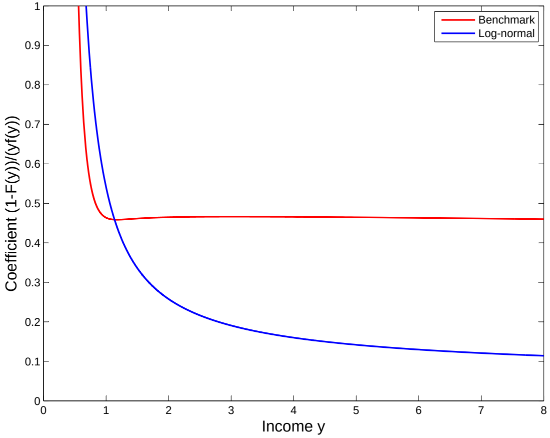
Sensitivity: Log-Normal Wage

- Log-normal distribution ⇒ thin right tail
- Optimal HSV worse than optimal affine
- Optimal affine nearly efficient

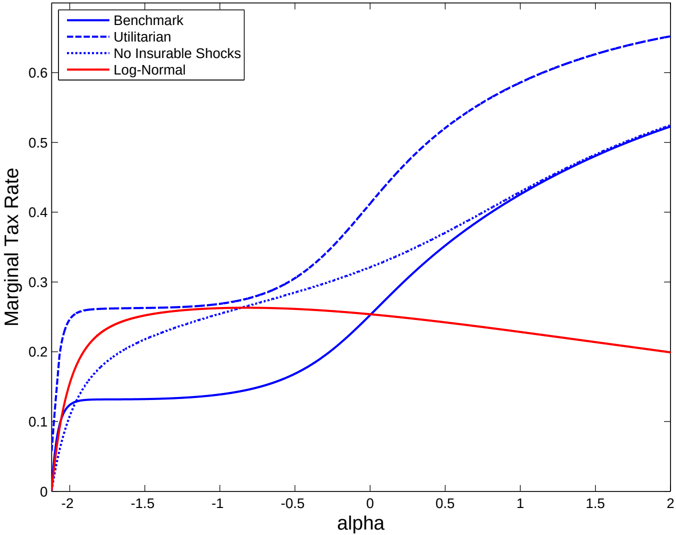
Tax System	Tax Parameters		Outcomes			
			welfare	Y	mar. tax	TR/Y
HSV^{US}	$\lambda : 0.836$	$\tau : 0.151$	–	–	0.311	0.018
HSV	$\lambda : 0.826$	$\tau : 0.070$	0.27	3.10	0.239	-0.005
Affine	$\tau_0 : -0.068$	$\tau_1 : 0.250$	0.34	2.67	0.250	0.042
Mirrlees			0.35	2.71	0.249	0.042

Why Distribution Shape Matters

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



Efficient Marginal Tax Rates: Sensitivity

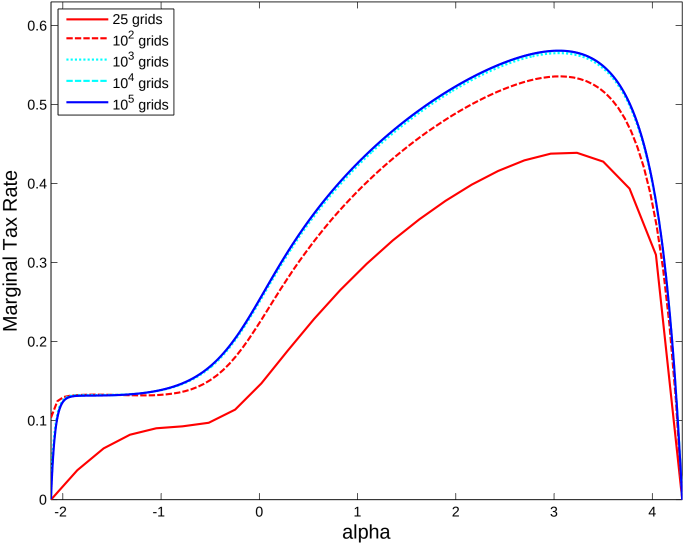


Extension: Coarse Grid

- Coarse grid $\Rightarrow$ Mirrlees Planner can do much better
- Gives Mirrlees planner too much power if true distribution continuous

# of grid points	Welfare (relative to HSV)		
	Affine	Mirrlees	First Best
25	-0.58	3.82	8.80
100	-0.58	1.17	8.79
1,000	-0.58	0.21	8.80
10,000	-0.58	0.11	8.80
100,000	-0.58	0.10	8.80

Extension: Coarse Grid



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 $\Rightarrow$ marginal tax rates of 20%

System		Outcomes			
		wel.	Y	mar.	TR/Y
HSV^{US}	$\lambda : 0.827, \tau : 0.151$	–	–	0.311	0.017 0.015
HSV	λ^L: 0.988, τ^L: 0.180 λ^H: 0.694, τ^H : −0.059	1.34	4.64	0.212	0.043 −0.061
Affine	τ_0^L: −0.140, τ_1^L: 0.151 τ_0^H: 0.095, τ_1^H : 0.224	1.39	5.20	0.199	0.126 −0.137
Mirrlees		1.46	5.20	0.200	0.103 −0.136

Conclusions

- Moving from current HSV system to optimal affine system is welfare reducing
  - Increasing marginal rates more important than lump-sum transfers
- Moving from current HSV system to fully optimal system generates tiny welfare gains
  - Ramsey and Mirrlees tax schemes not far apart: can approximately decentralize SB with a simple tax scheme
  - Important to measure the gap in terms of allocations and welfare, not in terms of marginal tax rates
- Want to condition both transfers and tax rates on observables