Optimal Income Taxation: Mirrlees Meets Ramsey

Jonathan Heathcote
FRB Minneapolis and CEPR

Hitoshi Tsujiyama
Goethe University Frankfurt

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)

- 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
   - \( \tau \) 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

- Agents differ wrt unobservable log productivity $\alpha$
- Planner only observes earnings $x = \exp(\alpha) \times h$
- Think of planner choosing $(c, x)$ for each $\alpha$ 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
     ⇒ 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 ⇒ Very progressive taxes
Environment 1

- Standard static Mirrlees plus partial private insurance (quantitatively important)

- Heterogeneous individual labor productivity with two stochastic components

  \[ \log w = \alpha + \varepsilon \]

- \( \varepsilon \) is privately-insurable, \( \alpha \) is not

  - Agents belong to large families
  - \( \alpha \) common across all members of a family \( \Rightarrow \) cannot be pooled within family
  - \( \varepsilon \) purely idiosyncratic & orthogonal to \( \alpha \) \( \Rightarrow \) 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 $\{c(\tilde{\alpha}), y(\tilde{\alpha})\}$

• Family heads draw idiosyncratic $\alpha$ and report $\tilde{\alpha}$

Second Stage

• Family members draw idiosyncratic $\varepsilon$

• Family head tells each member how much to work

• Total earnings must deliver $y(\tilde{\alpha})$ to the planner

• Must divide consumption $c(\tilde{\alpha})$ between family members
Nature of the Solution

- Planner cannot condition individual allocations on $\varepsilon$, 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
  - 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

\[ \max_{c(\alpha), y(\alpha)} \int W(\alpha) U(\alpha, \alpha) dF_\alpha \]

s.t. \[ \int y(\alpha) dF_\alpha \geq \int c(\alpha) dF_\alpha + G \]

\[ U(\alpha, \alpha) \geq U(\alpha, \tilde{\alpha}) \quad \forall \alpha, \forall \tilde{\alpha} \]

where \( U(\alpha, \tilde{\alpha}) \equiv \)

\[ \left\{ \begin{array}{l}
\max_{\{c(\alpha, \tilde{\alpha}, \varepsilon), h(\alpha, \tilde{\alpha}, \varepsilon)\}} \int \left\{ \log(c(\alpha, \tilde{\alpha}, \varepsilon)) - \frac{h(\alpha, \tilde{\alpha}, \varepsilon)^{1+\sigma}}{1+\sigma} \right\} dF_\varepsilon \\
\int c(\alpha, \tilde{\alpha}, \varepsilon) dF_\varepsilon = c(\tilde{\alpha}) \\
\int \exp(\alpha + \varepsilon) h(\alpha, \tilde{\alpha}, \varepsilon) dF_\varepsilon = y(\tilde{\alpha})
\end{array} \right. \]

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

where \( \Omega = \left( \int \exp(\varepsilon)^{\frac{1+\sigma}{\sigma}} dF_\varepsilon(\varepsilon) \right)^{-\sigma} \)
Planner’s Problem: Ramsey

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

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

\[
\begin{align*}
\max \{ &c(\alpha, \varepsilon), h(\alpha, \varepsilon)\} \quad \int \left\{ \log c(\alpha, \varepsilon) - \frac{h(\alpha, \varepsilon)^{1+\sigma}}{1+\sigma} \right\} dF_\varepsilon \\
\text{s.t.} \quad &\int c(\alpha, \varepsilon)dF_\varepsilon = y(\alpha) - T(y(\alpha); \tau) \\
&y(\alpha) = \int \exp(\alpha + \varepsilon)h(\alpha, \varepsilon)dF_\varepsilon
\end{align*}
\]
Social Preferences

- Assume SWF takes the form $W(\alpha; \theta) = \exp(-\theta \alpha)$
  - $\theta$ 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 $\theta$ gives observed $\tau$ as solution to Ramsey problem?

- Let $\tau^*(\theta)$ denote welfare-maximizing choice for $\tau$ given $\theta$

- 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 $\tau^{US}$
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
- $F_\alpha$: Exponentially Modified Gaussian $EMG(\mu_\alpha, \sigma^2_\alpha, \lambda_\alpha)$
- $F_\varepsilon$: Normal $N\left(-\frac{\sigma^2_\varepsilon}{2}, \sigma^2_\varepsilon\right)$
- $\log(w) = \alpha + \varepsilon$ 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: $\sigma^2_y = \left(\frac{1+\sigma}{\sigma+\tau}\right)^2 \sigma^2_\varepsilon + \sigma^2_\alpha$. 
Use micro data from the 2007 SCF to estimate $\alpha$ 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 $G \text{ s.t. } G/Y = 0.188$ (US, 2005)

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

  - 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

<table>
<thead>
<tr>
<th>Percentile Ratios</th>
<th>Model</th>
<th>LP</th>
</tr>
</thead>
<tbody>
<tr>
<td>P5/P1</td>
<td>1.48</td>
<td>1.48</td>
</tr>
<tr>
<td>P10/P5</td>
<td>1.24</td>
<td>1.20</td>
</tr>
<tr>
<td>P25/P10</td>
<td>1.44</td>
<td>1.40</td>
</tr>
</tbody>
</table>
Numerical Implementation

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

$\text{Density vs. Wage (exp(\alpha + \varepsilon))}$

$\times 10^{-3}$
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

<table>
<thead>
<tr>
<th>Tax System</th>
<th>Tax Parameters</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>welfare</td>
</tr>
<tr>
<td>HSV&lt;sup&gt;US&lt;/sup&gt;</td>
<td>$\lambda: 0.839$</td>
<td>$\tau: 0.161$</td>
</tr>
<tr>
<td>HSV</td>
<td>$\lambda: 0.817$</td>
<td>$\tau: 0.330$</td>
</tr>
<tr>
<td>Affine</td>
<td>$\tau_0: -0.259$</td>
<td>$\tau_1: 0.492$</td>
</tr>
<tr>
<td>Mirrlees</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Benchmark: Mirrlees vs Ramsey

A. Log Consumption

B. Hours Worked

C. Marginal Tax Rate

D. Average Tax Rate
Optimal HSV better than optimal affine

⇒ 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: $\tilde{v}_\alpha = v_\alpha + v_\epsilon$ and $\tilde{v}_\epsilon = 0$

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

<table>
<thead>
<tr>
<th>Tax System</th>
<th>Tax Parameters</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>welfare</td>
</tr>
<tr>
<td>HSV&lt;sub&gt;US&lt;/sub&gt;</td>
<td>$\lambda : 0.842$</td>
<td>$\tau : 0.161$</td>
</tr>
<tr>
<td>HSV</td>
<td>$\lambda : 0.804$</td>
<td>$\tau : 0.383$</td>
</tr>
<tr>
<td>Affine</td>
<td>$\tau_0 : -0.283$</td>
<td>$\tau_1 : 0.545$</td>
</tr>
<tr>
<td>Mirrlees</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 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 \( \theta^* \):
    \[
    \sigma^2_\alpha \theta^* - \frac{1}{\lambda_\alpha + \theta^*} = -\frac{1}{\lambda_\alpha - 1 + \tau} - \sigma^2_\alpha (1 - \tau) + \frac{1}{1+\sigma} \left\{ \frac{1}{(1-g)(1-\tau)} - 1 \right\}
    \]

- Simple in Normal case \( (\lambda_\alpha \to \infty) \)
  \[
  \theta^* = -(1 - \tau) + \frac{1}{\sigma^2_\alpha} \frac{1}{1 + \sigma} \left\{ \frac{1}{(1-g)(1-\tau)} - 1 \right\}
  \]
  - \( \theta^* \) increasing in \( \tau \) and \( g \)
  - \( \theta^* \) declining in \( \sigma \) and \( \sigma^2_\alpha \)
  - \( \theta^* \) increasing in \( \lambda_\alpha \) (holding fixed \( var(\alpha) = \sigma^2_\alpha + \frac{1}{\lambda_\alpha^2} \))
Social Welfare Functions

Relative Pareto Weight ($\exp(-\theta \alpha)$)

- **Laissez-Faire**: $\theta = -1$
- **Empirically Motivated**: $\theta^* = -0.566$
- **Utilitarian**: $\theta = 0$
### Sensitivity: Alternative SWFs

<table>
<thead>
<tr>
<th>SWF</th>
<th>( \theta )</th>
<th>( T'(y) )</th>
<th>( TR/Y )</th>
<th>( \Delta Y )</th>
<th>Mirrlees Allocations</th>
<th>Welfare Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laissez-Faire</td>
<td>-1</td>
<td>0.083</td>
<td>-0.082</td>
<td>9.72</td>
<td>3.15</td>
<td>3.14 2.98</td>
</tr>
<tr>
<td>Emp. Motivated</td>
<td>-0.57</td>
<td>0.314</td>
<td>0.051</td>
<td>0.16</td>
<td>0.05</td>
<td>-0.48</td>
</tr>
<tr>
<td>Utilitarian</td>
<td>0</td>
<td>0.491</td>
<td>0.213</td>
<td>-7.99</td>
<td>2.48</td>
<td>1.77 2.08</td>
</tr>
<tr>
<td>Rawlsian</td>
<td>( \infty )</td>
<td>0.711</td>
<td>0.538</td>
<td>-22.55</td>
<td>708.28</td>
<td>649.14 354.90</td>
</tr>
</tbody>
</table>
Empirically-Motivated SWF

A. Log Consumption

B. Hours Worked

C. Marginal Tax Rate

D. Average Tax Rate
HSV vs Affine with Various SWFs

Welfare Gains (%)

Taste for Redistribution ($\theta$)

Mirrlees
HSV
Affine

Welfare Gains (%)

Taste for Redistribution ($\theta$)
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 $G$

A. Log Consumption

B. Hours Worked

C. Marginal Tax Rate (with $\alpha$)

D. Marginal Tax Rate (with income)

- $g = 0.50$
- $g = 0.75$

Baseline $g = 0.50$. $g = 0.75$. Income ($y$).
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): $\frac{1+\sigma}{\sigma+\lambda}\approx 71\%$

- 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

A. Marginal Tax Rate (with $\alpha$)

- Baseline
- $\theta = 1$
- No private insurance
- Elastic labor, $\sigma = .5$

B. Marginal Tax Rate (with income)
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

<table>
<thead>
<tr>
<th>Tax System</th>
<th>Tax Parameters</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>welfare</td>
</tr>
<tr>
<td>HSV&lt;sub&gt;US&lt;/sub&gt;</td>
<td>$\lambda : 0.828$</td>
<td>$\tau : 0.161$</td>
</tr>
<tr>
<td>HSV</td>
<td>$\lambda : 0.813$</td>
<td>$\tau : 0.285$</td>
</tr>
<tr>
<td>Affine</td>
<td>$\tau_0 : -0.230$</td>
<td>$\tau_1 : 0.451$</td>
</tr>
<tr>
<td>Mirrlees</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 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

<table>
<thead>
<tr>
<th>Tax System</th>
<th>Tax Parameters</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSV$^{US}$</td>
<td>$\lambda$ 0.839 $\tau$ 0.161</td>
<td>welfare $-$ $-$ $0.319$ $0.018$</td>
</tr>
<tr>
<td>Affine</td>
<td>$\tau_0$ $-$0.259 $\tau_1$ 0.492</td>
<td>$1.77$ $-$8.00 $0.492$ $0.279$</td>
</tr>
<tr>
<td>Cubic</td>
<td>$\tau_0$ $-$0.212 $\tau_1$ 0.370 $\tau_2$ 0.049 $\tau_3$ $-$0.002</td>
<td>$2.40$ $-$8.01 $0.491$ $0.228$</td>
</tr>
<tr>
<td>Mirrlees</td>
<td>$-$2.248 $-$7.99 $0.491$ $0.213$</td>
<td></td>
</tr>
</tbody>
</table>
Cubic Tax Function

A. Log Consumption

B. Hours Worked

C. Marginal Tax Rate

D. Average Tax Rate
Extension: Type-Contingent Taxes

- Productivity partially reflects observable characteristics (e.g. education, age, gender)

- Some fraction of uninsurable shocks are observable: 
  \[ \alpha \rightarrow \alpha + \kappa \]

- Heathcote, Perri & Violante (2010) estimate variance of cross-sectional wage dispersion attributable to observables, \( \nu_\kappa = 0.108 \)

- Planner should condition taxes on observables: \( T(y; \kappa) \)

- Consider two-point distribution for \( \kappa \) (college vs high school)
### Extension: Type-Contingent Taxes

- **Significant welfare gains** relative to non-contingent tax
- **Conditioning on observables** $\Rightarrow$ marginal tax rates of 42%

<table>
<thead>
<tr>
<th>System</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wel.</td>
</tr>
<tr>
<td>HSV&lt;sub&gt;US&lt;/sub&gt;</td>
<td>$\lambda : 0.834, \tau : 0.161$</td>
</tr>
<tr>
<td>HSV</td>
<td>$\lambda^L : 1.069, \tau^L : 0.480$</td>
</tr>
<tr>
<td>HSV</td>
<td>$\lambda^H : 0.595, \tau^H : 0.073$</td>
</tr>
<tr>
<td>Affine</td>
<td>$\tau^L_0 : -0.403, \tau^L_1 : 0.345$</td>
</tr>
<tr>
<td>Affine</td>
<td>$\tau^H_0 : -0.032, \tau^H_1 : 0.452$</td>
</tr>
<tr>
<td>Mirrlees</td>
<td>$-$</td>
</tr>
<tr>
<td></td>
<td>$-$</td>
</tr>
</tbody>
</table>
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