Optimal Income Taxation: Mirrlees Meets Ramsey

Jonathan Heathcote
FRB Minneapolis

Hitoshi Tsujiyama
Goethe University Frankfurt

Research in Economic Dynamics Group, Sep 2013
How should we tax income?

- Want to maximize social insurance while minimizing distortions to labor supply
- Tension between the two objectives:
  - Taxes highly distorting ⇒ costly to redistribute ⇒ want low taxes and transfers (US model)
  - Large gains from redistribution ⇒ want high taxes and transfers (European model)
Approaches to Tax Design

• Mirrlees
  • Solve for constrained efficient allocations

• Ramsey
  • Ad hoc restrictions on the shape of the tax schedule
  • But can explore tax design in richer quantitative models

• Our Goals
  1. Extend Mirrlees approach to provide quantitative guidance on tax design
  2. Explore approximate decentralizations of constrained efficient allocations in Ramsey tradition
Contributions

• Quantitative challenges for the Mirrlees approach

  1. Planner provides all insurance in the economy  
     ⇒ decentralizations feature highly progressive taxes
  2. No observable component of productivity  
     ⇒ exaggerates information friction
  3. Ad hoc choice for social welfare function  
     ⇒ prescriptions likely sensitive to this

• Our responses

  1. Explicitly model partial private insurance
  2. Introduce observable differences in productivity alongside non-observable differences
  3. Construct a social welfare function motivated by the observed tax system
Preview of Findings

- Observed tax system suggests US social welfare function over-weights high productivity workers

- Large potential welfare gains from moving to constrained efficient tax system

- Simple affine tax schedule delivers most of these gains, ...

- ... but important to condition both transfers and tax rates on observables
Environment

- Preferences
  \[ U(c, h) = \frac{c^{1-\gamma}}{1-\gamma} - \frac{h^{1+\sigma}}{1+\sigma} \]

- Production
  \[ Y = \int whdi = \int cdi + G \]

- Individual labor productivity
  \[ \log w = \alpha + \kappa + \varepsilon \]

- \( \varepsilon \): perfect private insurance
- \( \alpha \) and \( \kappa \): only social insurance through tax system
- \( \alpha \): private information
- \( \kappa \): reflects observables (e.g. age, education, gender)
Planner’s Problems

1. First best: planner observes $\alpha$ and $\kappa$

2. Second best (Mirrlees):
   $\alpha$ is private, unrestricted non-linear taxes

3. Third best (Ramsey):
   $\alpha$ is private, restricted tax functions
   - Polynomial non-type-specific functions of earnings or income
   - $\kappa$-type-specific tax functions

   - Planner always takes private insurance for $\varepsilon$ as given.
   - Social welfare function $W(\alpha, \kappa)$
Planner’s Problem: First/Second Best

Timeline:

1. First Stage
   - Draw \((\alpha, \kappa)\), report \(\tilde{\alpha}\)
   - Make a contract with the planner

2. Second Stage
   - Buy insurance against the insurable shock \(\varepsilon\), draw \(\varepsilon\)
   - Choose labor supply and receive insurance payments
   - Consume \(c(\tilde{\alpha}, \kappa)\), deliver \(y(\tilde{\alpha}, \kappa)\) to the planner
Planner’s Problem: Second Best

\[
\max_{c(\alpha, \kappa), y(\alpha, \kappa)} \int \int W(\alpha, \kappa) U(\alpha, \alpha, \kappa) dF_\alpha dF_\kappa \\
\text{s.t.} \quad \int \int y(\alpha, \kappa) dF_\alpha dF_\kappa \geq \int \int c(\alpha, \kappa) dF_\alpha dF_\kappa + G \\
U(\alpha, \alpha, \kappa) \geq U(\tilde{\alpha}, \alpha, \kappa) \quad \forall (\alpha, \kappa), \forall \tilde{\alpha}
\]

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

\[
\begin{cases}
\max_{h, B} \int \left( \frac{c(\tilde{\alpha}, \kappa)^{1-\gamma}}{1-\gamma} - \frac{h(\varepsilon; \tilde{\alpha}, \alpha, \kappa)^{1+\sigma}}{1+\sigma} \right) dF_\varepsilon \\
\text{s.t.} \quad \int Q(\varepsilon) B(\varepsilon; \tilde{\alpha}, \alpha, \kappa) d\varepsilon = 0 \\
\exp(\alpha + \kappa + \varepsilon) h(\varepsilon; \tilde{\alpha}, \alpha, \kappa) + B(\varepsilon; \tilde{\alpha}, \alpha, \kappa) = y(\tilde{\alpha}, \kappa) \quad \forall \varepsilon
\end{cases}
\]

Price of insurance \( Q = f(\varepsilon) \)
Planner’s Problem: Second Best

- Solving the second stage problem and substituting its solution into the first stage,

\[
\max_{c(\alpha,\kappa),y(\alpha,\kappa)} \int \int W(\alpha, \kappa) \left( \frac{c(\alpha,\kappa)^{1-\gamma}}{1-\gamma} - \frac{\Omega}{1+\sigma} \left( \frac{y(\alpha,\kappa)}{\exp(\alpha+\kappa)} \right)^{1+\sigma} \right) dF_\alpha dF_\kappa
\]

subject to

\[
\int \int y(\alpha, \kappa) dF_\alpha dF_\kappa \geq \int \int c(\alpha, \kappa) dF_\alpha dF_\kappa + G
\]

\[
\frac{c(\alpha,\kappa)^{1-\gamma}}{1-\gamma} - \frac{\Omega}{1+\sigma} \left( \frac{y(\alpha,\kappa)}{\exp(\alpha+\kappa)} \right)^{1+\sigma} \geq \frac{c(\tilde{\alpha},\kappa)^{1-\gamma}}{1-\gamma} - \frac{\Omega}{1+\sigma} \left( \frac{y(\tilde{\alpha},\kappa)}{\exp(\alpha+\kappa)} \right)^{1+\sigma}
\]

where \( \Omega \equiv \left( \int \exp(\varepsilon) \frac{1+\sigma}{\sigma} dF_\varepsilon \right)^{-\sigma} \)

- This is a standard Mirrlees Problem. Can be decentralized income (or consumption) taxes.
Planner’s Problem: Ramsey

- Consider income taxes $T(y(\alpha, \kappa, \varepsilon))$
- Ramsey planner chooses $T$ to maximize the social welfare given agents’ problem:

$$\max_{c(\alpha,\kappa,\varepsilon), h(\alpha,\kappa,\varepsilon), B(\alpha,\kappa,\varepsilon)} \int \left( \frac{c(\alpha,\kappa,\varepsilon)^{1-\gamma}}{1-\gamma} - \frac{h(\alpha,\kappa,\varepsilon)^{1+\sigma}}{1+\sigma} \right) dF_{\varepsilon}$$

subject to

$$\int Q(\varepsilon)B(\alpha, \kappa, \varepsilon)d\varepsilon = 0$$

$$c(\alpha, \kappa, \varepsilon) \leq y(\alpha, \kappa, \varepsilon) - T(y(\alpha, \kappa, \varepsilon)) \quad \forall \varepsilon$$

where

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

- Under a mild assumption on $T$, can show that $c$ and $y$ are independent of $\varepsilon$
- Also consider earnings taxation:

$$T(x(\alpha, \kappa, \varepsilon)) \text{ where } x(\alpha, \kappa, \varepsilon) \equiv \exp(\alpha + \kappa + \varepsilon)h(\alpha, \kappa, \varepsilon)$$
Baseline Parameterization

- Log consumption: $\gamma = 1$
- Frisch elasticity $= 0.5 \rightarrow \sigma = 2$
- Gov C + Gov I = 18.8% of GDP in 2005 $\rightarrow g = 0.188$

- $v_\varepsilon$: estimated var of insurable shocks
  $\rightarrow v_\varepsilon = 0.193$ (Heathcote, Storesletten & Violante (2013))

- $v_\kappa$: estimated var of cross-sectional wage dispersion attributable to observables
  $\rightarrow v_\kappa = 0.108$ (Heathcote, Perri & Violante (2010))

- Total variance of wages is 0.466 (Heathcote, Storesletten & Violante (2011))
  $\rightarrow v_\alpha = 0.165$
Wage Distributions

• 2 point equal-weight distribution for $\kappa$:
  \[ \exp(\kappa_H)/\exp(\kappa_L) = 1.93 \]

• Log-normal distribution for $\varepsilon$ (continuous)

• Bounds for $\alpha$:
  \[ \exp(\alpha) \in \left[ \frac{1}{2} \times 5.15, \frac{200.56}{19.60} \right] \]
  - $19.60$ is average hourly earnings in 2005 (BLS CES)
  - $5.15$ is Federal minimum wage in 2005
  - $200.56$ is earnings per hour at 99.5$^{th}$ percentile of 2005 earnings distribution (Piketty & Saez, Table B3)

• Distribution: Log-normal for $\exp(\alpha) \leq x$, Pareto for $\exp(\alpha) > x$
  - $x = 1.77$ (95% in log-normal range, Pareto above $\$34.6$)
  - $a = 2.0$ (Pareto parameter estimated from Piketty & Saez)

• Grid for $\alpha$ : 10,000 evenly spaced points
Baseline Tax System

- Tax / transfer system applies to earnings $x = wh$
- Schedule given by
  \[ T(x) = x - \lambda x^{1-\tau} \]
  - Functional form introduced by Benabou (2000)
  - Closely approximate US tax transfer system (Heathcote, Storesletten, and Violante, 2013)
  - $\tau$ indexes progressivity of the tax system
Fit of HSV Tax function
Social Welfare

- Infer social preferences from the degree of progressivity built into the actual tax/transfer system
- Assume SWF takes the form
  \[ W(\alpha, \kappa) = \exp(-\theta (\alpha + \kappa)) \]
- Assume US govt is choosing a tax/transfer system in the class
  \[ T(x) = x - \lambda x^{1-\tau} \]
- What value for \( \theta \) best rationalizes the observed choice for \( \tau \)?
Social Welfare

- Let $\tau^*(\theta)$ be welfare maximizing $\tau$ given $\theta$.
- *Empirically-motivated social welfare function: $\theta^{US}$ that solves*

$$
\tau^*(\theta^{US}) = \tau^{US}
$$

- Special case: if $\gamma = 1$ and $F$’s are log-normal,

$$
\theta^{US} = \tau^{US} - 1 + \frac{1}{\nu_\alpha + \nu_\kappa} \left\{ \frac{1}{1-g} \left[ \frac{1}{(1+\sigma)(1-\tau^{US})} + \frac{\tau^{US}(1+\sigma)^2}{(\sigma+\tau^{US})^3} \nu_\varepsilon \right] - \frac{1}{1+\sigma} \right\}
$$

- In general, can solve for $\theta^{US}$ numerically
- Use $\theta^{US}$ as baseline for welfare comparisons

$\Rightarrow$ Welfare improving tax reforms reflect improved efficiency, rather than choice for SWF
Social Welfare

![Graph showing the relationship between Hourly Wages and Relative Pareto Weight. The graph depicts a curve that increases as Hourly Wages increase.]
1. Mirrlees:
   - Solve exactly on a discrete grid by forward iteration
   - Much faster and more accurate than looking for approximate solution to Saez’ functional equation

2. Ramsey:
   Use Mirrlees allocations, run regression to find first guess for tax system
Computation

- Take HSV earnings taxation $T(x) = x - \lambda x^{1-\tau}$ as the benchmark.
- Relative to this evaluate:
  1. Mitrless second best
  2. Wide class of optimal income tax functions:
     - $\tau^0 + \tau^1 y$
     - $... + \tau^2 y^2$
     - $... + \tau^3 y^3$
  3. Experiment with / without $\kappa$–type specific coefficients
Result: No Type-Contingent taxes

<table>
<thead>
<tr>
<th>Tax system</th>
<th>HSV</th>
<th>$\lambda$</th>
<th>$\tau$</th>
<th>welfare</th>
<th>$Y$</th>
<th>mar. tax</th>
<th>$G/Y$</th>
<th>TR/Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>earnings</td>
<td>0.852</td>
<td>0.180</td>
<td>-</td>
<td>-</td>
<td>0.293</td>
<td>0.188</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>income</td>
<td>0.828</td>
<td>0.215</td>
<td>0.43</td>
<td>-1.17</td>
<td>0.276</td>
<td>0.190</td>
<td>0.036</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\tau^0$</th>
<th>$\tau^1$</th>
<th>$\tau^2$</th>
<th>$\tau^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>0.175</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-0.179</td>
<td>0.377</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-0.162</td>
<td>0.340</td>
<td>0.012</td>
<td>-</td>
</tr>
<tr>
<td>-0.139</td>
<td>0.284</td>
<td>0.041</td>
<td>-0.002</td>
</tr>
</tbody>
</table>

Second Best (Mirrlees) | 4.31 | 3.57 | 0.288 | 0.182 | 0.264 |
First Best             | 18.22 | 20.77 | 0     | 0.156 | 0.510 |
Result: No Type-Contingent taxes

- Relative to HSV tax scheme, Ramsey schemes imply output losses (1.5%)

- Optimal marginal tax rates around 30%

- Large transfers for the low income agents

- Lump-sum component key to welfare gains
  - quadratic term adds small
  - cubic term generates further gains
## Result: Type-Contingent Taxes

<table>
<thead>
<tr>
<th></th>
<th>( \lambda )</th>
<th>( \tau )</th>
<th>welfare</th>
<th>( Y )</th>
<th>mar. tax</th>
<th>TR/Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>earnings</td>
<td>0.852</td>
<td>0.180</td>
<td>-</td>
<td>-</td>
<td>0.293</td>
<td>0.026</td>
</tr>
<tr>
<td>income</td>
<td>1.067</td>
<td>0.368</td>
<td>3.69</td>
<td>2.15</td>
<td>0.351</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>0.625</td>
<td>-0.007</td>
<td></td>
<td></td>
<td></td>
<td>-0.072</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \tau^0 )</th>
<th>( \tau^1 )</th>
<th>( \tau^2 )</th>
<th>( \tau^3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.139</td>
<td>0.284</td>
<td>0.041</td>
<td>-0.002</td>
</tr>
<tr>
<td>-0.130</td>
<td>-0.334</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-0.343</td>
<td>0.304</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>0.098</td>
<td>0.080</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-0.123</td>
<td>0.241</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-0.314</td>
<td>0.332</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>0.066</td>
<td>0.052</td>
<td>0.133</td>
<td>-0.012</td>
</tr>
<tr>
<td>-0.259</td>
<td>0.146</td>
<td>0.062</td>
<td>-0.003</td>
</tr>
</tbody>
</table>

Second Best (Mirrlees) 4.31 3.57 0.288 0.264
Result: Type-Contingent Taxes

• Significant welfare gains relative to non-contingent tax

• Affine tax schedule delivers 88.5% of the welfare gain of SB

• Want both \( \tau^0 \) and \( \tau^1 \) to be \( \kappa \)-type specific:
  
  • Redistribute across \( \kappa \)-types via \( \tau_H^1 > \tau_L^1 \)
  
  • If only linear taxes are available, set \( \frac{1-\tau_H^1}{1-\tau_L^1} = \frac{\exp(\kappa_L)}{\exp(\kappa_H)} \)
  
  • Induce higher hours from \( \kappa_H \)-type via \( \tau_H^0 > \tau_L^0 \)

• Approximately implement SB with type-specific cubic tax (96%)
Baseline: HSV Tax

Graphs showing consumption, hours worked, marginal tax rate, and net tax in response to log hourly wages for different scenarios of tax rates (Low and High) under Mirrlees and Ramsey frameworks.
Type-dependent HSV Tax

Consumption vs. Log Hourly Wages
- Low (Mirrlees)
- High (Mirrlees)
- Low (Ramsey)
- High (Ramsey)

Hours Worked vs. Log Hourly Wages

Marginal Tax Rate vs. Log Hourly Wages

Net Tax vs. Log Hourly Wages
Affine Tax

- Consumption
- Log Hourly Wages (Low (Mirrlees), High (Mirrlees), Low (Ramsey), High (Ramsey))
- Hours Worked
- Log Hourly Wages
- Marginal Tax Rate
- Net Tax
- Log Hourly Wages
Sensitivity 1: Number of Grid Points for $\alpha$

<table>
<thead>
<tr>
<th># of grid points</th>
<th>Ramsey Affine</th>
<th>Ramsey Cubic</th>
<th>Mirrlees</th>
<th>First Best</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>3.63</td>
<td>4.11</td>
<td>5.77</td>
<td>18.18</td>
</tr>
<tr>
<td>1,000</td>
<td>3.66</td>
<td>4.14</td>
<td>4.45</td>
<td>18.21</td>
</tr>
<tr>
<td>10,000</td>
<td>3.66</td>
<td>4.14</td>
<td>4.30</td>
<td>18.22</td>
</tr>
</tbody>
</table>
Coarseness of Grid

Marginal Tax Rate vs. Log Hourly Wages

- Low (10^4)
- High (10^4)
- Low (10^3)
- High (10^3)
Sensitivity 1: Number of Grid Points for $\alpha$

- Coarse grid $\Rightarrow$ Mirrlees Planner can do much better
- Linear taxes no longer approximate the Second Best
- Coarse grids give Mirrlees planner too much power if true distribution is continuous
### Sensitivity Analysis 2: Social Welfare Function

**Utilitarian Social Welfare Function:** \( W(\alpha, \kappa) = 1 \)

<table>
<thead>
<tr>
<th>Alternative Welfare Functions: Ramsey Affine vs Mirrlees Second Best</th>
<th>Ramsey Affine</th>
<th>Welfare</th>
<th>mar. tax</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau^0 )</td>
<td>( \tau^1 )</td>
<td>Rmsy</td>
<td>SB</td>
<td>Rmsy</td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td>-0.314</td>
<td>0.241</td>
<td>3.66</td>
<td>4.31</td>
</tr>
<tr>
<td></td>
<td>0.066</td>
<td>0.332</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Utilitarian</strong></td>
<td>-0.398</td>
<td>0.276</td>
<td>6.03</td>
<td>6.76</td>
</tr>
<tr>
<td></td>
<td>0.072</td>
<td>0.376</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sensitivity Analysis 2: Social Welfare Function

- With Utilitarian SWF, the optimal tax displays larger transfers for the low type and higher marginal rates
- Lower output gains, higher welfare gains for Utilitarian case
- The optimal tax structure is sensitive to the choice of SWF
Conclusions

1. Debate on the structure of labor income taxation: how to balance redistribution and distortion welfare gains

2. Ramsey and Mirrlees tax schemes are not very far apart: can approximately decentralize Mirrlees solution with a very simple tax scheme

⇒ Important to measure the gap in terms of allocations and welfare, not in terms of marginal tax rates

3. Want to condition both transfers and tax rates on observables

4. Working on extension to Pareto improving tax reform