Pareto Efficient Income Taxation

Iván Werning
MIT

April 2007
NBER Public Economics meeting
Q: Good shape for tax schedule?
Q: Good shape for tax schedule?

  - positive: redistribution vs. efficiency
  - normative: Utilitarian social welfare function
Q: Good shape for tax schedule?

  - positive: redistribution vs. efficiency
  - normative: Utilitarian social welfare function

- this paper: Pareto efficient taxation
  - positive: redistribution vs. efficiency
  - normative: Utilitarian social welfare function

Pareto Efficiency
Why not Utilitarian? \( \left( \sum_i U^i \right) \)

- **practical**: cardinality \( U^i \rightarrow W(U^i) \) (or even \( W^i(U^i) \))
  ... which Utilitarian?

- **conceptual**: political process:
  social classes \( \rightarrow \) Coasian bargain
  ... but \( \max \sum U^i \) ?

- **philosophical**: other notions of fairness and social justice
Old Motivation: “New New New New...”

- Why not Utilitarian? ($\sum U^i$)
  - practical: cardinality $U^i \rightarrow W(U^i)$ (or even $W^i(U^i)$)
    ... which Utilitarian?
  - conceptual: political process:
    social classes $\rightarrow$ Coasian bargain
    ...but $\max \sum U^i$?
  - philosophical: other notions of fairness and social justice

- Pareto efficiency $\rightarrow$ weaker criterion
Pareto Frontier

\[ V_L \]

\[ V_H \]

first best

constrained

Pareto Efficient Income Taxation
Pareto Frontier

V_L

V_H
Pareto Frontier
Pareto Frontier

V_L

V_H

Pareto Efficient Income Taxation
Pareto Frontier
Contribution

- invert Mirrlees model...
- ...express in tractable way
- ...use it: some applications
Results

#0 restrictions generalize “zero-tax-at-the-top”

#1 Any \( T(Y) \)…

▷ efficient for many \( f(\theta) \)

▷ inefficient for many \( f(\theta) \)

… anything goes

#2 Given \( T_0(Y) \) \( \rightarrow \) \( g(Y) \) \( \rightarrow \) \( f(\theta) \) (Saez, 2001)

▷ efficient set of \( T(Y) \): large

▷ inefficient set of \( T(Y) \): large

#3 Simple test for efficiency of \( T_0(Y) \)
#4 Simple formulas...
- bound on top tax rate
- efficiency of a flat tax

#5 Increasing progressivity
- maintains Pareto efficiency

#6 observable heterogeneity
- not conditioning can be efficient
Positive side of Mirrlees (1971)

- continuum of types $\theta \sim F(\theta)$

- additive preferences

\[ U(c, Y, \theta) = u(c) - \theta h(Y) \]

(e.g. $Y = w \cdot n$ and $h(n) = \alpha n^n$)
Positive side of Mirrlees (1971)

- continuum of types $\theta \sim F(\theta)$

- additive preferences

\[ U(c, Y, \theta) = u(c) - \theta h(Y) \]

(e.g. $Y = w \cdot n$ and $h(n) = \alpha n^n$)

- given $T(Y)$

\[ v(\theta) \equiv \max_Y U(Y - T(Y), Y, \theta) \]
Positive side of Mirrlees (1971)

- continuum of types $\theta \sim F(\theta)$

- additive preferences

\[ U(c, Y, \theta) = u(c) - \theta h(Y) \]

(e.g. $Y = w \cdot n$ and $h(n) = \alpha n^\eta$)

- given $T(Y)$

\[ v(\theta) \equiv \max_Y U(Y - T(Y), Y, \theta) \]

- Government budget

\[ \int T(Y(\theta)) \, dF(\theta) \geq G \]
Positive side of Mirrlees (1971)

- continuum of types $\theta \sim F(\theta)$
- additive preferences
  \[ U(c, Y, \theta) = u(c) - \theta h(Y) \]
  (e.g. $Y = w \cdot n$ and $h(n) = \alpha n^n$)
- given $T(Y)$
  \[ v(\theta) \equiv \max_{Y} U(Y - T(Y), Y, \theta) \]
- Resource feasible
  \[ \int (Y(\theta) - c(\theta)) \, dF(\theta) \geq G \]
Positive side of Mirrlees (1971)

- continuum of types \( \theta \sim F(\theta) \)
- additive preferences
  \[
  U(c, Y, \theta) = u(c) - \theta h(Y)
  \]
  (e.g. \( Y = w \cdot n \) and \( h(n) = \alpha n^\eta \))
- given \( T(Y) \)
  \[
  u'(\theta) = U_\theta(Y(\theta) - T(Y(\theta)), Y(\theta), \theta)
  \]
- Resource feasible
  \[
  \int (Y(\theta) - c(\theta)) \, dF(\theta) \geq G
  \]
Positive side of Mirrlees (1971)

- continuum of types \( \theta \sim F(\theta) \)
- additive preferences
  \[
  U(c, Y, \theta) = u(c) - \theta h(Y)
  \]
  (e.g. \( Y = w \cdot n \) and \( h(n) = \alpha n^\eta \))
- given \( T(Y) \)
  \[
  v'(\theta) = -h(Y(\theta))
  \]
- Resource feasible
  \[
  \int (Y(\theta) - c(\theta)) \, dF(\theta) \geq G
  \]
Setup

Positive side of Mirrlees (1971)

- continuum of types $\theta \sim F(\theta)$
- additive preferences
  \[ U(c, Y, \theta) = u(c) - \theta h(Y) \]
  (e.g. $Y = w \cdot n$ and $h(n) = \alpha n^n$)
- given $T(Y)$
  \[ v'(\theta) = -h(Y(\theta)) \]
- Resource feasible
  \[ \int (Y(\theta) - e(v(\theta), Y(\theta), \theta)) \, dF(\theta) \geq G \]
Planning Problem

Dual Pareto Problem

maximize net resources

subject to,

\[ \tilde{v}(\theta) \geq v(\theta) \]

incentives
Dual Pareto Problem

\[
\max_{\tilde{Y}, \tilde{v}} \int (\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)) \, dF(\theta)
\]

subject to,

\[
\tilde{v}(\theta) \geq v(\theta)
\]

incentives
Planning Problem

Dual Pareto Problem

\[
\max_{\tilde{Y}, \tilde{v}} \int (\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)) dF(\theta)
\]

subject to,

\[
\tilde{v}(\theta) \geq v(\theta)
\]

\[
\tilde{v}'(\theta) = -h(\tilde{Y}(\theta))
\]
Dual Pareto Problem

\[
\max_{\tilde{Y}, \tilde{v}} \int (\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)) \, dF(\theta)
\]

subject to,

\[
\tilde{v}(\theta) \geq v(\theta)
\]

\[
\tilde{v}'(\theta) = -h(\tilde{Y}(\theta))
\]

\(\tilde{Y}(\theta)\) nonincreasing
Efficiency Conditions

Lagrangian

\[
\mathcal{L} = \int \left( \tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta) \right) dF(\theta) \\
- \int \left( \tilde{v}'(\theta) + h(\tilde{Y}(\theta)) \right) \mu(\theta) d\theta
\]
Efficiency Conditions

Lagrangian (integrating by parts)

\[ \mathcal{L} = \int \left( \tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta) \right) dF(\theta) - \tilde{v}(\bar{\theta})\mu(\bar{\theta}) + \mu(\theta)\tilde{v}(\theta) + \int \tilde{v}(\theta)\mu'(\theta)d\theta - \int h(\tilde{Y}(\theta))\mu(\theta) d\theta \]
Efficiency Conditions

Lagrangian (integrating by parts)

\[ \mathcal{L} = \int (\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)) \, dF(\theta) - \tilde{v}(\tilde{\theta})\mu(\tilde{\theta}) + \mu(\theta)\tilde{v}(\theta) \]

\[ + \int \tilde{v}(\theta)\mu'(\theta)d\theta - \int h(\tilde{Y}(\theta))\mu(\theta) \, d\theta \]

First-order conditions

\[ (1 - e_Y(v(\theta), Y(\theta), \theta))f(\theta) = \mu(\theta)h'(Y(\theta)) \quad [Y(\theta)] \]
Efficiency Conditions

Lagrangian (integrating by parts)

\[ \mathcal{L} = \int (\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)) \, dF(\theta) - \tilde{v}(\bar{\theta})\mu(\bar{\theta}) + \mu(\theta)\tilde{v}(\theta) \]
\[ + \int \tilde{v}(\theta)\mu'(\theta) \, d\theta - \int h(\tilde{Y}(\theta))\mu(\theta) \, d\theta \]

First-order conditions

\[ \tau(\theta)f(\theta) = \mu(\theta)h'(Y(\theta)) \]
Lagrangian (integrating by parts)

\[ \mathcal{L} = \int \left( \tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta) \right) dF(\theta) - \tilde{v}(\bar{\theta}) \mu(\bar{\theta}) + \mu(\theta) \tilde{v}(\theta) + \int \tilde{v}(\theta) \mu'(\theta) d\theta - \int h(\tilde{Y}(\theta)) \mu(\theta) d\theta \]

First-order conditions

\[ \mu(\theta) = \tau(\theta) \frac{f(\theta)}{h'(Y(\theta))} \]
Efficiency Conditions

Lagrangian (integrating by parts)

\[ L = \int \left( \tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta) \right) dF(\theta) - \tilde{v}(\theta)\mu(\theta) + \mu(\theta)\tilde{v}(\theta) \]

\[ + \int \tilde{v}(\theta)\mu'(\theta)d\theta - \int h(\tilde{Y}(\theta))\mu(\theta) d\theta \]

First-order conditions

\[ \mu(\theta) = \tau(\theta) \frac{f(\theta)}{h'(Y(\theta))} \]

\[ \mu'(\theta) \leq e_v(v(\theta), Y(\theta), \theta) f(\theta) \]
Efficiency Conditions

Lagrangian (integrating by parts)

\[
\mathcal{L} = \int \left( \tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta) \right) dF(\theta) - \tilde{v}(\bar{\theta})\mu(\bar{\theta}) + \mu(\theta)\tilde{v}(\theta) \\
+ \int \tilde{v}(\theta)\mu'(\theta)d\theta - \int h(\tilde{Y}(\theta))\mu(\theta) d\theta
\]

First-order conditions

\[
\mu(\theta) = \tau(\theta) \frac{f(\theta)}{h(Y(\theta))} \quad [Y(\theta)] \\
\mu'(\theta) \leq e_v(v(\theta), Y(\theta), \theta) f(\theta) \quad [v(\theta)]
\]

\[
\tau(\theta) \left( \theta \frac{\tau'(\theta)}{\tau(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 1 - \tau(\theta)
\]
Efficiency Conditions

Proposition. \( T(Y) \) is Pareto efficient if and only if

\[
\tau(\theta) \left( \theta \frac{\tau'(\theta)}{\tau(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 1 - \tau(\theta)
\]

\[
\tau(\bar{\theta}) \geq 0 \quad \text{and} \quad \tau(\theta) \leq 0.
\]
Proposition. \( T(Y) \) is Pareto efficient if and only if

\[
\tau(\theta) \left( \frac{\theta}{\tau(\theta)} \frac{\tau'(\theta)}{\tau(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 1 - \tau(\theta)
\]

\[
\tau(\bar{\theta}) \geq 0 \quad \text{and} \quad \tau(\theta) \leq 0.
\]

Note: “zero-tax-at-top” special case
Efficiency Conditions

Proposition. \( T(Y) \) is Pareto efficient if and only if

\[
\tau(\theta) \left( \frac{\tau'(\theta)}{\tau(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 1 - \tau(\theta)
\]

\[
\tau(\bar{\theta}) \geq 0 \quad \text{and} \quad \tau(\theta) \leq 0.
\]

- note: “zero-tax-at-top” special case
- more general condition:

\[
\frac{\tau(\theta)f(\theta)}{h'(Y(\theta))} + \int_{\theta}^{\bar{\theta}} \frac{1}{u'(c(\tilde{\theta}))} f(\tilde{\theta}) \ d\tilde{\theta}
\]

is nonincreasing.
Intuition

Define

\[ \hat{T}(Y) \equiv \begin{cases} T(Y(\hat{\theta})) - \varepsilon & Y = Y(\hat{\theta}) \\ T(Y) & Y \neq Y(\hat{\theta}) \end{cases} \]

Proposition. \( \hat{T} \succeq T \leftrightarrow \)

\[
\tau(\theta) \left( \theta \frac{\tau'(\theta)}{\tau(\theta)} + 2 \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 3(1 - \tau(\theta))
\]

is violated at \( \hat{\theta} \)
Simple Tax Reform

- Introduction
- Model
- Main Results
  - Intuition
  - Anything Goes
  - Identification and Test
  - Graphical Test
  - Empirical Strategy
  - Quantifying Inefficiencies
- Applications
- Conclusions
Simple Tax Reform

Pareto Efficient Income Taxation
Simple Tax Reform

Introduction

Model

Main Results

❖ Intuition
❖ Anything Goes
❖ Identification and Test
❖ Graphical Test
❖ Empirical Strategy
❖ Quantifying Inefficiencies

Applications

Conclusions

Pareto Efficient Income Taxation
Simple Tax Reform

Introduction

Model

Main Results
❖ Intuition
❖ Anything Goes
❖ Identification and Test
❖ Graphical Test
❖ Empirical Strategy
❖ Quantifying Inefficiencies

Applications

Conclusions

Pareto Efficient Income Taxation
Simple Tax Reform

Introduction
Model
Main Results
- Intuition
  - Anything Goes
  - Identification and Test
  - Graphical Test
  - Empirical Strategy
  - Quantifying Inefficiencies
Applications
Conclusions

Pareto Efficient Income Taxation
Simple Tax Reform

\[ \frac{g'(Y)}{g(Y)} \text{ small} \quad \left( \frac{f''(\theta)}{f(\theta)} \text{ large} \right) \rightarrow \text{inefficiency} \]
Laffer

- lower taxes → increase revenue
- Pareto improvements ↔ “Laffer” effect

Proposition. $T_1(Y) \geq T_0(Y) \rightarrow T_1(Y) \leq T_0(Y)$
\[ \tau(\theta) \left( \theta \frac{\tau'(\theta)}{\tau(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 1 - \tau(\theta) \]

**Proposition.** For any \( T(Y) \)
- exists set \( \{f(\theta)\} \rightarrow \text{Pareto efficient} \)
- exists set \( \{f(\theta)\} \rightarrow \text{Pareto inefficient} \)
\[ \tau(\theta) \left( \frac{\tau'(\theta)}{\tau(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 1 - \tau(\theta) \]

**Proposition.** For any $T(Y)$
- exists set $\{f(\theta)\}$ → Pareto efficient
- exists set $\{f(\theta)\}$ → Pareto inefficient

- without empirical knowledge
  → anything goes
Anything Goes

\[ \tau(\theta) \left( \theta \frac{\tau'(\theta)}{\tau(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(Y(\theta))}{d \log \theta} \right) \leq 1 - \tau(\theta) \]

Proposition. For any \( T(Y) \)

- exists set \( \{ f(\theta) \} \) \( \rightarrow \) Pareto efficient
- exists set \( \{ f(\theta) \} \) \( \rightarrow \) Pareto inefficient

- without empirical knowledge \( \rightarrow \) anything goes

- need information on \( f(\theta) \) to restrict \( T(Y) \)
observe $g(Y)$ identify (Saez, 2001)

$$
\theta(Y) = (1 - T'(Y)) \frac{u'(Y - T(Y))}{h'(Y)}
$$

$$
f(\theta(Y)) = \frac{g(Y)}{\theta'(Y)}
$$
observe \( g(Y) \) identify (Saez, 2001)

\[
\theta(Y) = (1 - T'(Y)) \frac{u'(Y - T(Y))}{h(Y)}
\]

\[ f(\theta(Y)) = \frac{g(Y)}{\theta'(Y)} \]

efficiency test...

\[
\frac{d \log g(Y)}{d \log Y} \geq a(Y)
\]

... for tax schedule in place
Graphical Test

- define Rawlsian density:

\[
\alpha(Y) = \frac{\exp\left(\int_0^Y a(z) \, dz\right)}{\int_0^\infty \exp\left(\int_0^Y a(z) \, dz\right)}
\]

- graphical test:

\[
\frac{g(Y)}{\alpha(Y)} \quad \text{nondecreasing}
\]
Empirical Implementation

needed

1. current tax function $T(Y)$
2. distribution of income $g(Y)$
3. utility function $U(c, Y, \theta)$
Intuition
Anything Goes
Identification and Test
Empirical Test
Empirical Strategy
Quantifying Inefficiencies

Empirical Implementation

- **needed**
  1. current tax function $T(Y)$
  2. distribution of income $g(Y)$
  3. utility function $U(c, Y, \theta)$

- in principle: #1 and #2 easy
  #3 usual deal
Empirical Implementation

- needed
  1. current tax function $T(Y)$
  2. distribution of income $g(Y)$
  3. utility function $U(c, Y, \theta)$

- in principle: #1 and #2 easy
  #3 usual deal

- Diamond (1998) and Saez (2001)
Empirical Implementation

- needed
  1. current tax function $T(Y)$
  2. distribution of income $g(Y)$
  3. utility function $U(c, Y, \theta)$

- in principle: #1 and #2 easy
  #3 usual deal

- Diamond (1998) and Saez (2001)

- some challenges...
  1. econometric: need to estimate $g'(Y)$ and $g(Y)$
  2. conceptual: static model
     lifetime $T(Y)$ and $g(Y)$ (Fullerton and Rogers)
Output Density

- IRS’s SOI Public Use Files for Individual tax returns
  - lifetime $g(Y)$?
  - lifetime $T(Y)$ schedule?

- $Y^i = \frac{1}{n} \sum Y^i_t$

- smooth density estimate assumed $T(Y) = .30 \times Y$
IRS’s SOI Public Use Files for Individual tax returns

- lifetime $g(Y)$?
- lifetime $T(Y)$ schedule?

$$Y^i = \frac{1}{n} \sum Y^i$$

smooth density estimate assumed $T(Y) = .30 \times Y$

**Figure 1:** Density of income $g(Y)$

**Figure 2:** Implied elasticity $Y g'(Y)$
IRS’s SOI Public Use Files for Individual tax returns

- lifetime \( g(Y) \)?
- lifetime \( T(Y) \) schedule?

\[ Y^i = \frac{1}{n} \sum Y^i_t \]

smooth density estimate assumed \( T(Y) = 0.30 \times Y \)
Quantifying Inefficiencies

- Efficiency test → qualitative
- Quantitative...

\[ \Delta \equiv \int (\tilde{Y}^*(\theta) - \tilde{c}^*(\theta)) \, dF(\theta) - \int (Y(\theta) - c(\theta)) \, dF(\theta) \]

- Does not count welfare improvements

\[ \tilde{v}(\theta) > v(\theta) \]
Top Tax Rate

- $u(c) = c^{1-\sigma}/(1 - \sigma)$ and $h(Y) = \alpha Y^\eta$

- Suppose top tax rate

\[ \bar{\tau} \equiv \lim_{\theta \to 0} \tau(\theta) = \lim_{Y \to \infty} T'(Y) \]

exists
Top Tax Rate

- $u(c) = c^{1-\sigma} / (1 - \sigma)$ and $h(Y) = \alpha Y^\eta$

- suppose top tax rate

$$\bar{\tau} \equiv \lim_{\theta \to 0} \tau(\theta) = \lim_{Y \to \infty} T'(Y)$$

exists

- efficiency condition bound

$$\bar{\tau} \leq \frac{\sigma + \eta - 1}{\varphi + \eta - 2}.$$ 

where $\varphi = -\lim_{T \to \infty} d \log g(Y) / d \log Y$. 

- Top Tax Rate
- Flat Tax
- Progressivity
- Heterogeneity

Conclusions
Top Tax Rate

- $u(c) = c^{1-\sigma}/(1 - \sigma)$ and $h(Y) = \alpha Y^\eta$

- suppose top tax rate

\[
\bar{\tau} \equiv \lim_{\theta \to 0} \tau(\theta) = \lim_{Y \to \infty} T'(Y)
\]

exists

- efficiency condition bound

\[
\bar{\tau} \leq \frac{\sigma + \eta - 1}{\varphi + \eta - 2}.
\]

where $\varphi = -\lim_{T \to \infty} d \log g(Y)/d \log Y$.

- Saez (2001): $\varphi = 3$
Flat Tax

- Linear tax necessary condition
  \[ \bar{\tau} \leq \frac{\sigma + \eta - 1}{-\frac{d\log g(Y)}{d\log Y} + \eta - 2} \]

- Linear tax sufficient condition
  \[ \bar{\tau} \leq \frac{\eta - 1}{-\frac{d\log g(Y)}{d\log Y} + \eta - 1} \]
Progressivity

- Quasi-linear $u(c) = c$

- result: can always increase progressivity
Heterogeneity

- groups = 1, ..., N

\[ f^i(\theta) \quad \text{and} \quad U^i(c, Y, \theta) \]

- unobservable \( i \)
  - single \( T(Y) \)
  - average efficiency condition

- observable \( i \)
  - multiple \( T^i(Y) \)
  - \( N \) efficiency conditions

- observation:
  - \( T^i(Y) = T(Y) \) may be Pareto efficient
  - never optimal for Utilitarian
## Conclusions

- Pareto efficiency → simple condition
- generalizes zero-tax-at-the-top result
- Pareto inefficient → Laffer effects
- flat taxes may be optimal...
- ...more progressivity always efficient