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 ?

- Mirrlees (1971), Diamond (1998), Saez (2001)
  - ▷ positive: redistribution vs. efficiency
  - ▷ normative: Utilitarian social welfare function

Q: Good shape for tax schedule ?

- Mirrlees (1971), Diamond (1998), Saez (2001)
  - 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

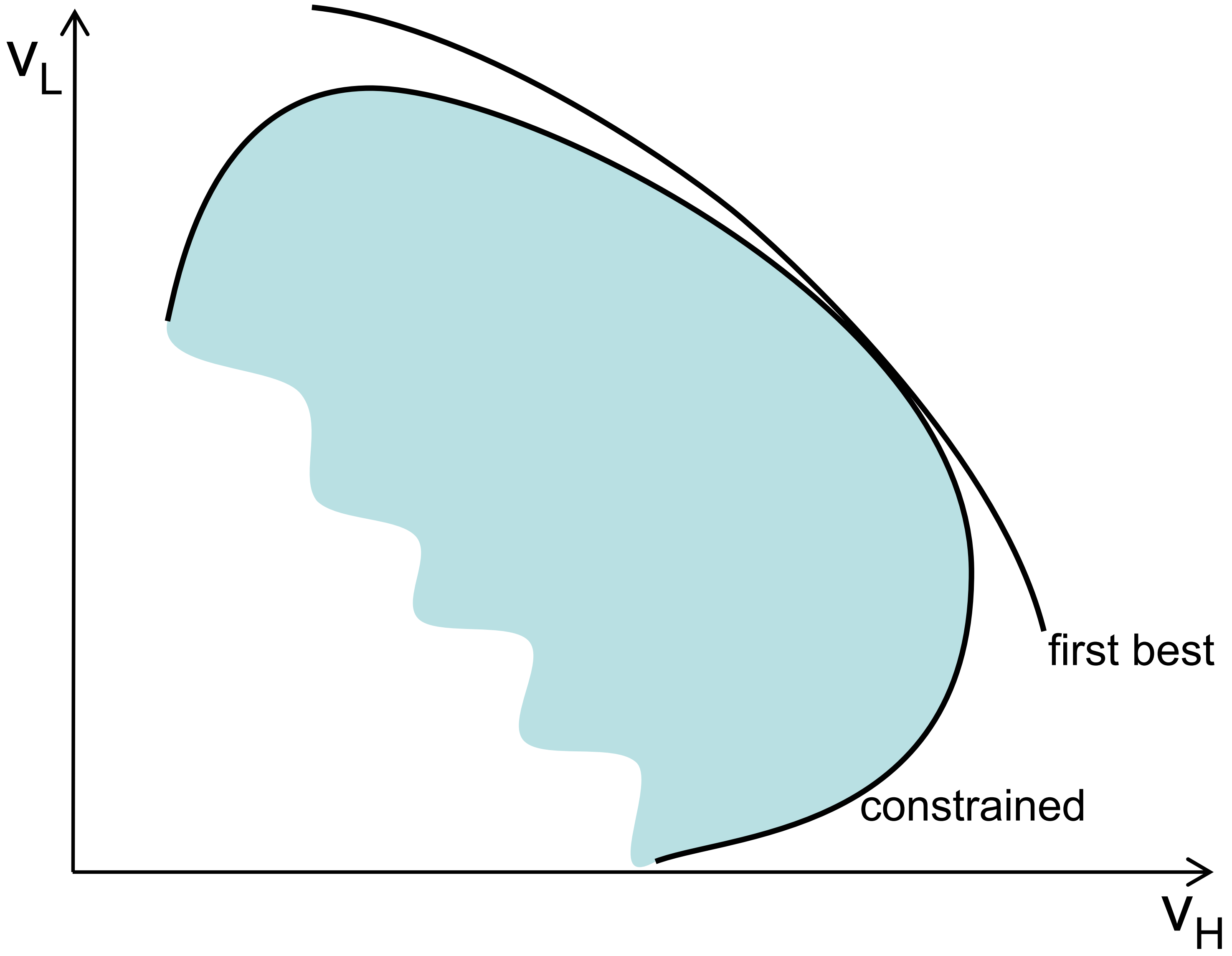
Old Motivation: "New New New..."

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

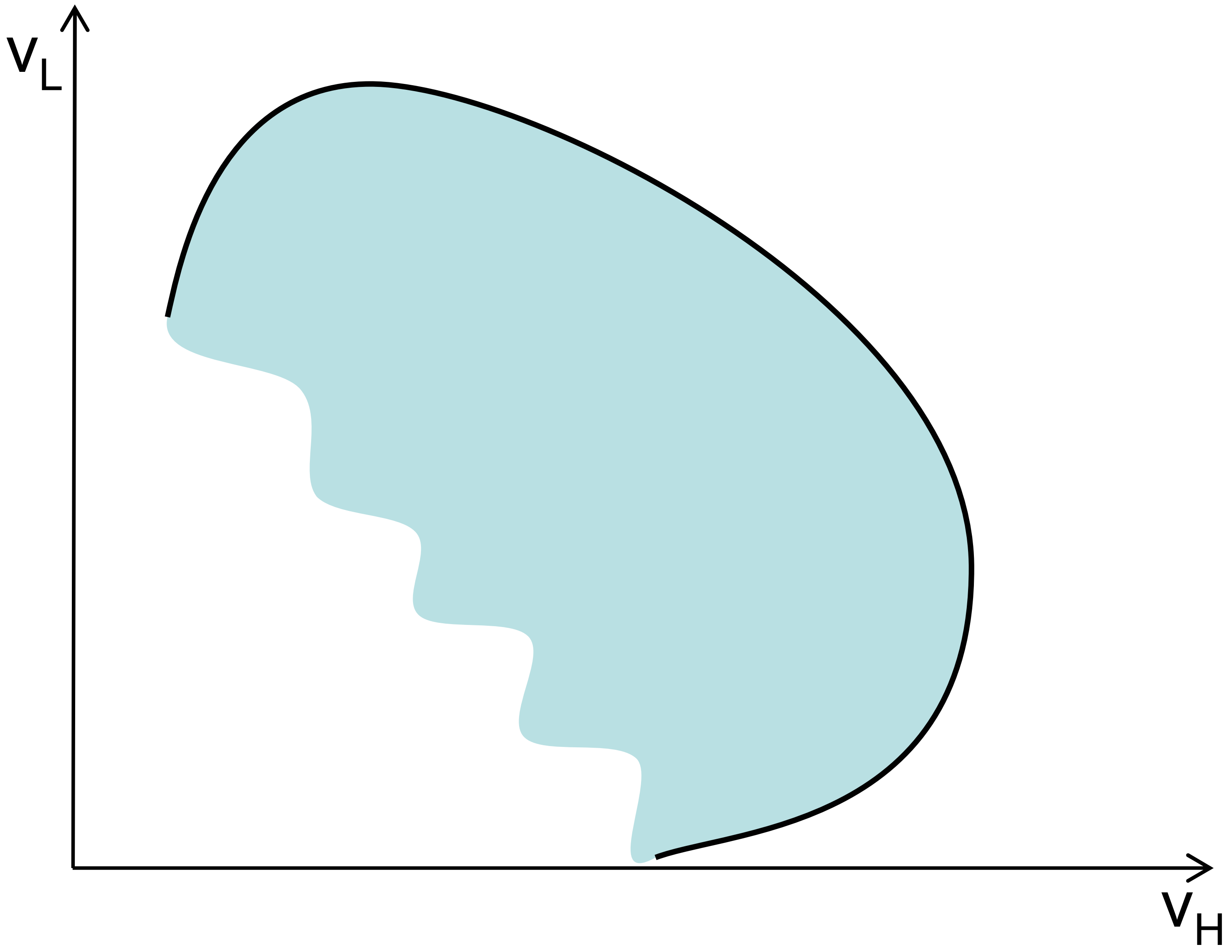
Old Motivation: "New New New..."

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

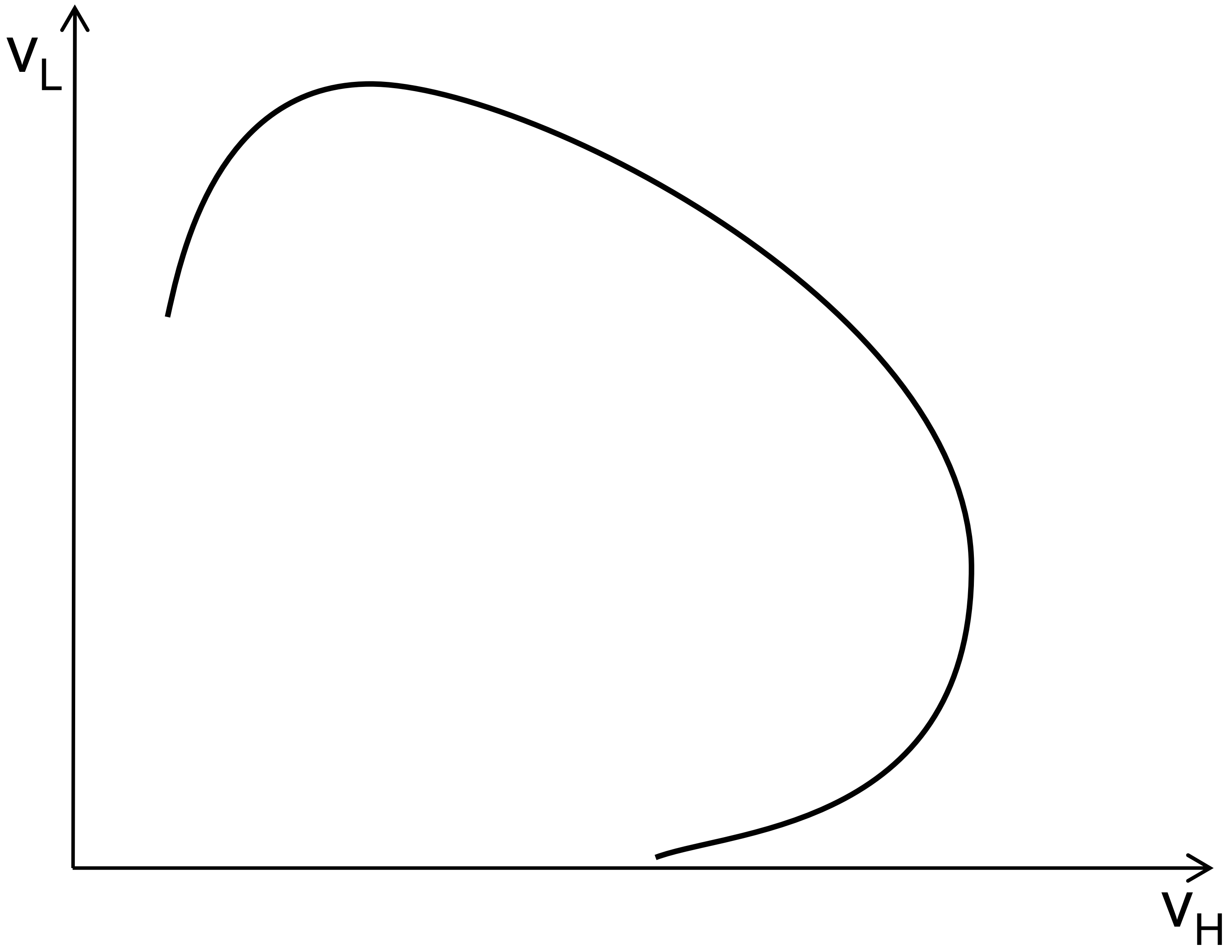
Pareto Frontier



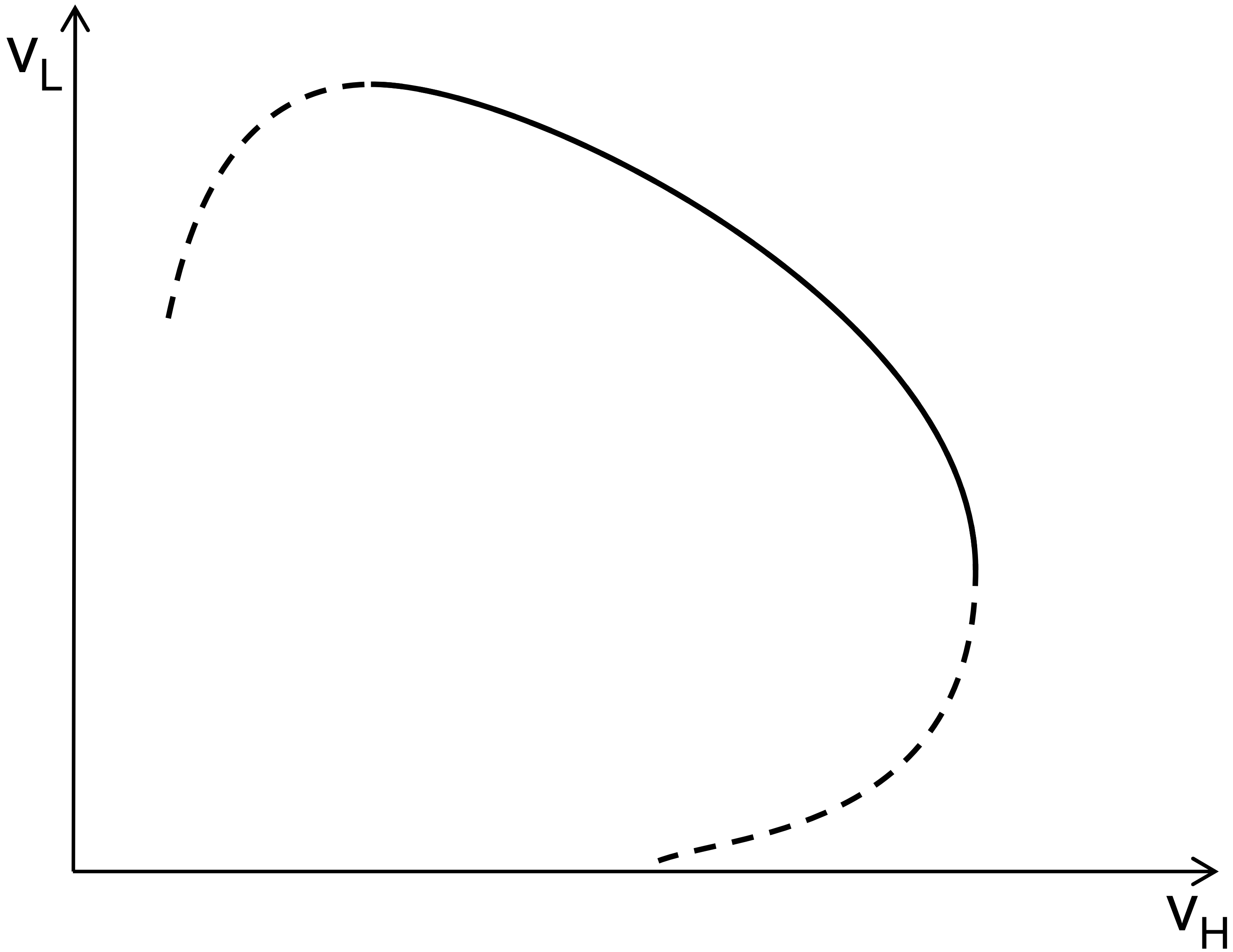
Pareto Frontier



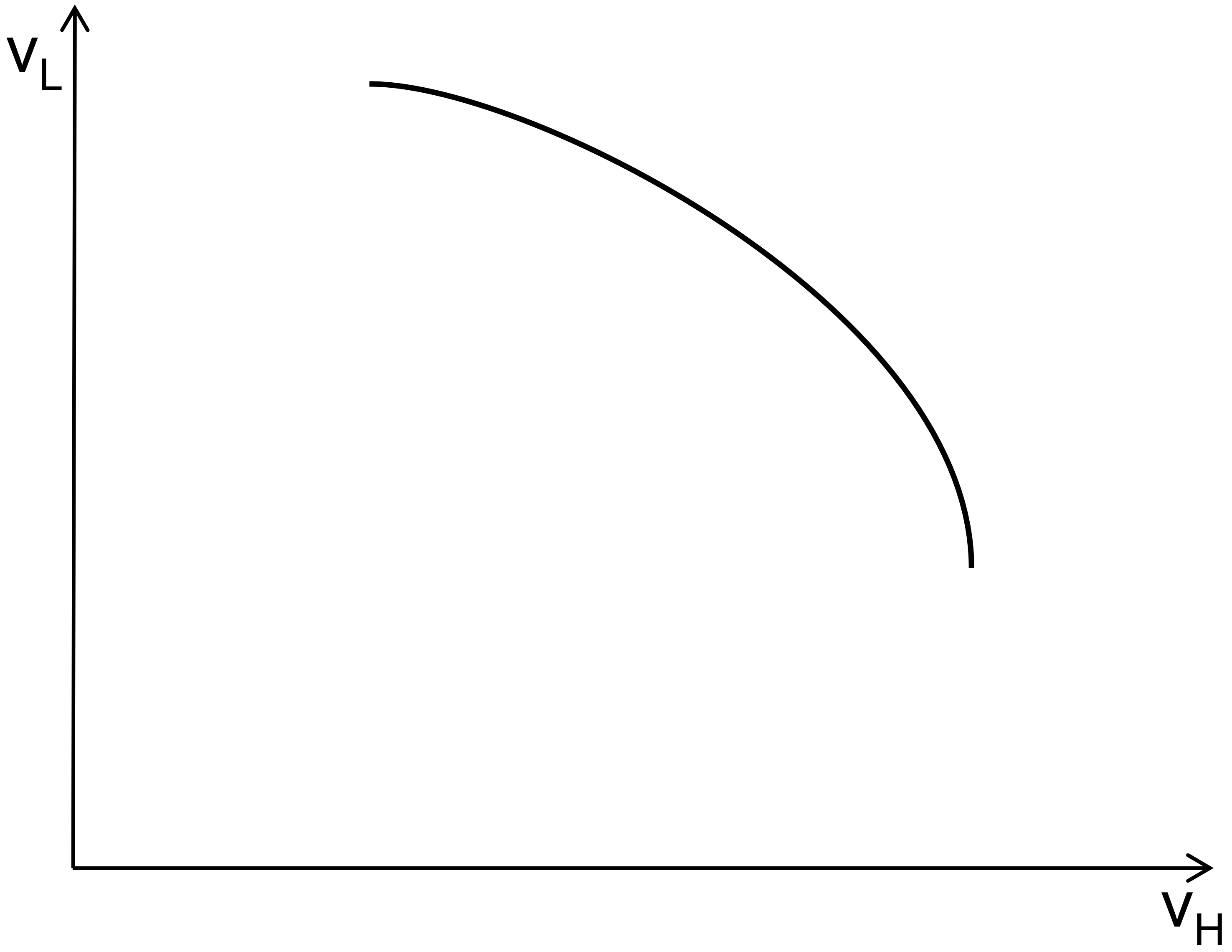
Pareto Frontier



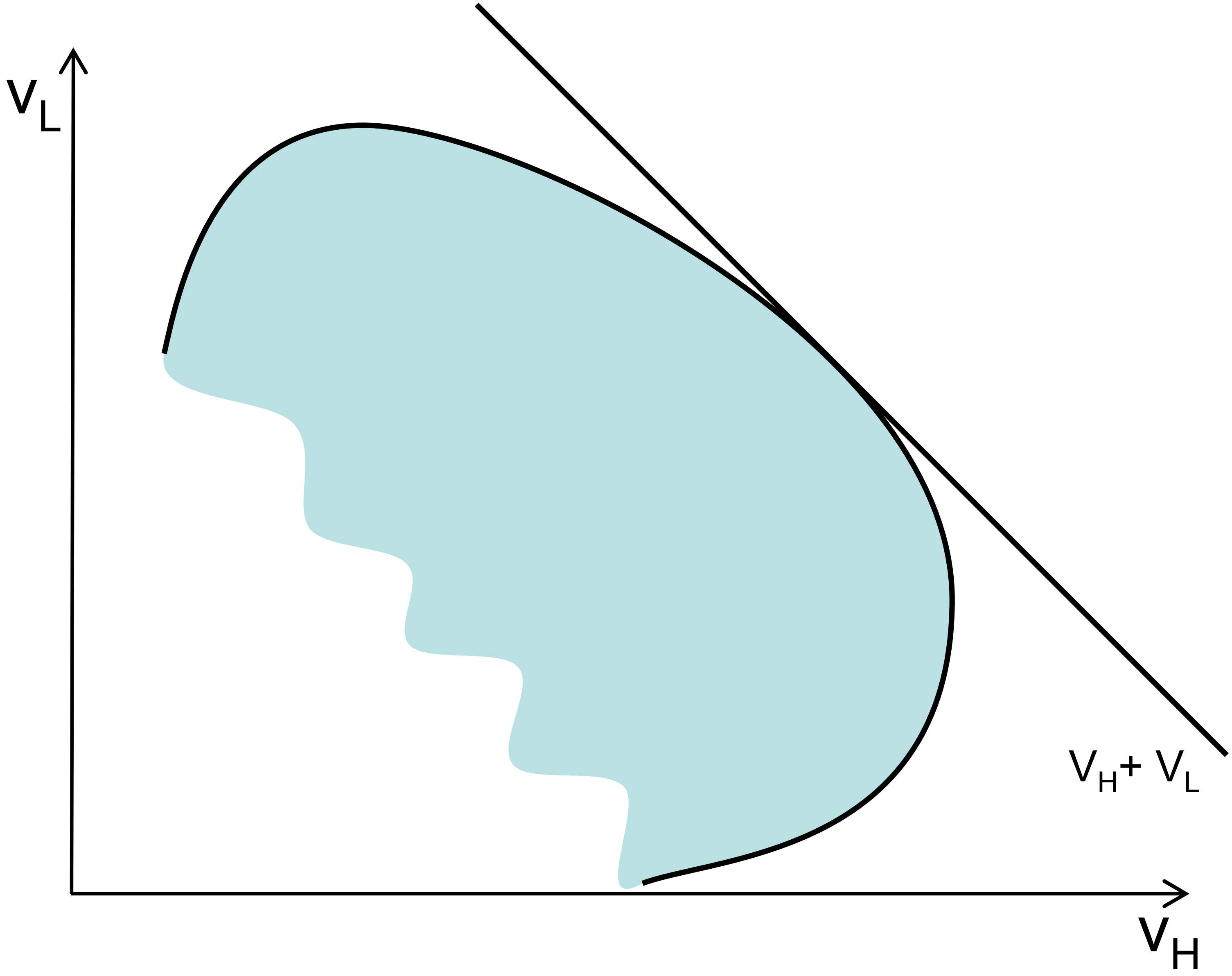
Pareto Frontier



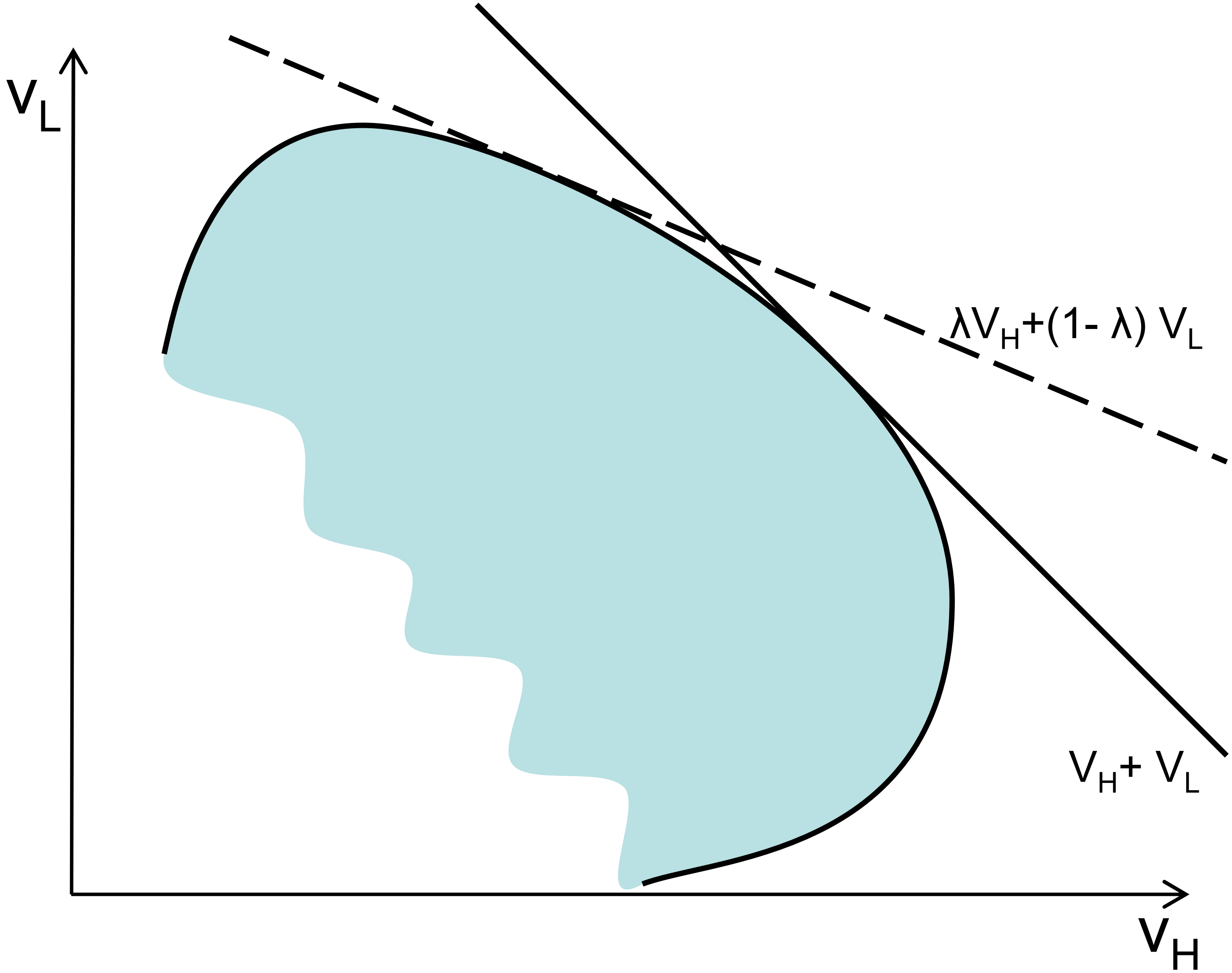
Pareto Frontier



Pareto Frontier



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 $T(Y)\ldots$

- ▷ efficient for many $f(\theta)$
- ▷ inefficient for many $f(\theta)$

$\ldots$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 $T_0(Y)$

Results

#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

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 $h(n) = \alpha n^\eta$)

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 $h(n) = \alpha n^{\eta}$)

- given $T(Y)$

$$v(\theta) \equiv \max_{Y} U(Y - T(Y), Y, \theta)$$

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 $h(n) = \alpha n^{\eta}$)

- given $T(Y)$

$$v(\theta) \equiv \max_{Y} U(Y - T(Y), Y, \theta)$$

- Government budget

$$\int T(Y(\theta))\, dF(\theta) \geq G$$

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 $h(n) = \alpha n^{\eta}$)

- given $T(Y)$

$$v(\theta) \equiv \max_{Y} U(Y - T(Y), Y, \theta)$$

- Resource feasible

$$\int \big(Y(\theta) - c(\theta)\big)\, dF(\theta) \geq G$$

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 $h(n) = \alpha n^{\eta}$)

- given $T(Y)$

$$v'(\theta) = U_{\theta}(Y(\theta) - T(Y(\theta)), Y(\theta), \theta)$$

- Resource feasible

$$\int \big(Y(\theta) - c(\theta)\big)\, dF(\theta) \geq G$$

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 $h(n) = \alpha n^{\eta}$)

- given $T(Y)$

$$v'(\theta) = -h(Y(\theta))$$

- Resource feasible

$$\int \big(Y(\theta) - c(\theta)\big)\, dF(\theta) \geq G$$

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

$$\int \left(Y(\theta) - e(v(\theta), Y(\theta), \theta)\right) dF(\theta) \geq G$$

Dual Pareto Problem

maximize net resources

subject to,

$$\tilde{v}(\theta) \geq v(\theta)$$

incentives

Dual Pareto Problem

$$\max_{\tilde{Y},\tilde{v}} \int \left(\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)\right) dF(\theta)$$

subject to,

$$\tilde{v}(\theta) \geq v(\theta)$$

incentives

Dual Pareto Problem

$$\max_{\tilde{Y},\tilde{v}} \int \left(\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)\right) dF(\theta)$$

subject to,

$$\tilde{v}(\theta) \geq v(\theta)$$

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

Dual Pareto Problem

$$\max_{\tilde{Y},\tilde{v}} \int \left(\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)\right) dF(\theta)$$

subject to,

$$\tilde{v}(\theta) \geq v(\theta)$$

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

$\tilde{Y}(\theta)$ nonincreasing

Efficiency Conditions

Lagrangian

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

Lagrangian (integrating by parts)

$$\mathcal{L} = \int \left(\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)\right) dF(\theta) - \tilde{v}(\bar{\theta})\mu(\bar{\theta}) + \mu(\underline{\theta})\tilde{v}(\underline{\theta})$$

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

Lagrangian (integrating by parts)

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

First-order conditions

$$\left(1 - e_Y(v(\theta), Y(\theta), \theta)\right) f(\theta) = \mu(\theta) h'(Y(\theta)) \qquad [Y(\theta)]$$

Efficiency Conditions

Lagrangian (integrating by parts)

$$\mathcal{L} = \int \left(\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)\right) dF(\theta) - \tilde{v}(\bar{\theta})\mu(\bar{\theta}) + \mu(\underline{\theta})\tilde{v}(\underline{\theta})$$

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

First-order conditions

$$\tau(\theta)f(\theta) = \mu(\theta)h'(Y(\theta)) \qquad [Y(\theta)]$$

Efficiency Conditions

Lagrangian (integrating by parts)

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

First-order conditions

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

Lagrangian (integrating by parts)

$$\mathcal{L} = \int \big(\tilde{Y}(\theta) - e(\tilde{v}(\theta), \tilde{Y}(\theta), \theta)\big)\, dF(\theta) - \tilde{v}(\bar{\theta})\mu(\bar{\theta}) + \mu(\underline{\theta})\tilde{v}(\underline{\theta})$$

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

First-order conditions

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

$$\mu'(\theta) \leq e_v\big(v(\theta), Y(\theta), \theta\big) f(\theta) \qquad [v(\theta)]$$

Efficiency Conditions

Lagrangian (integrating by parts)

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

First-order conditions

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

$$\mu'(\theta) \leq e_v\big(v(\theta), Y(\theta), \theta\big) f(\theta) \qquad [v(\theta)]$$

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

Efficiency Conditions

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

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

$\tau(\bar{\theta})\geq 0$ and $\tau(\underline{\theta})\leq 0.$

Efficiency Conditions

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

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

$\tau(\bar{\theta})\geq 0$ and $\tau(\underline{\theta})\leq 0.$

- note: “zero-tax-at-top” → special case

Efficiency Conditions

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

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

$\tau(\bar{\theta})\geq 0$ and $\tau(\underline{\theta})\leq 0.$

- 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

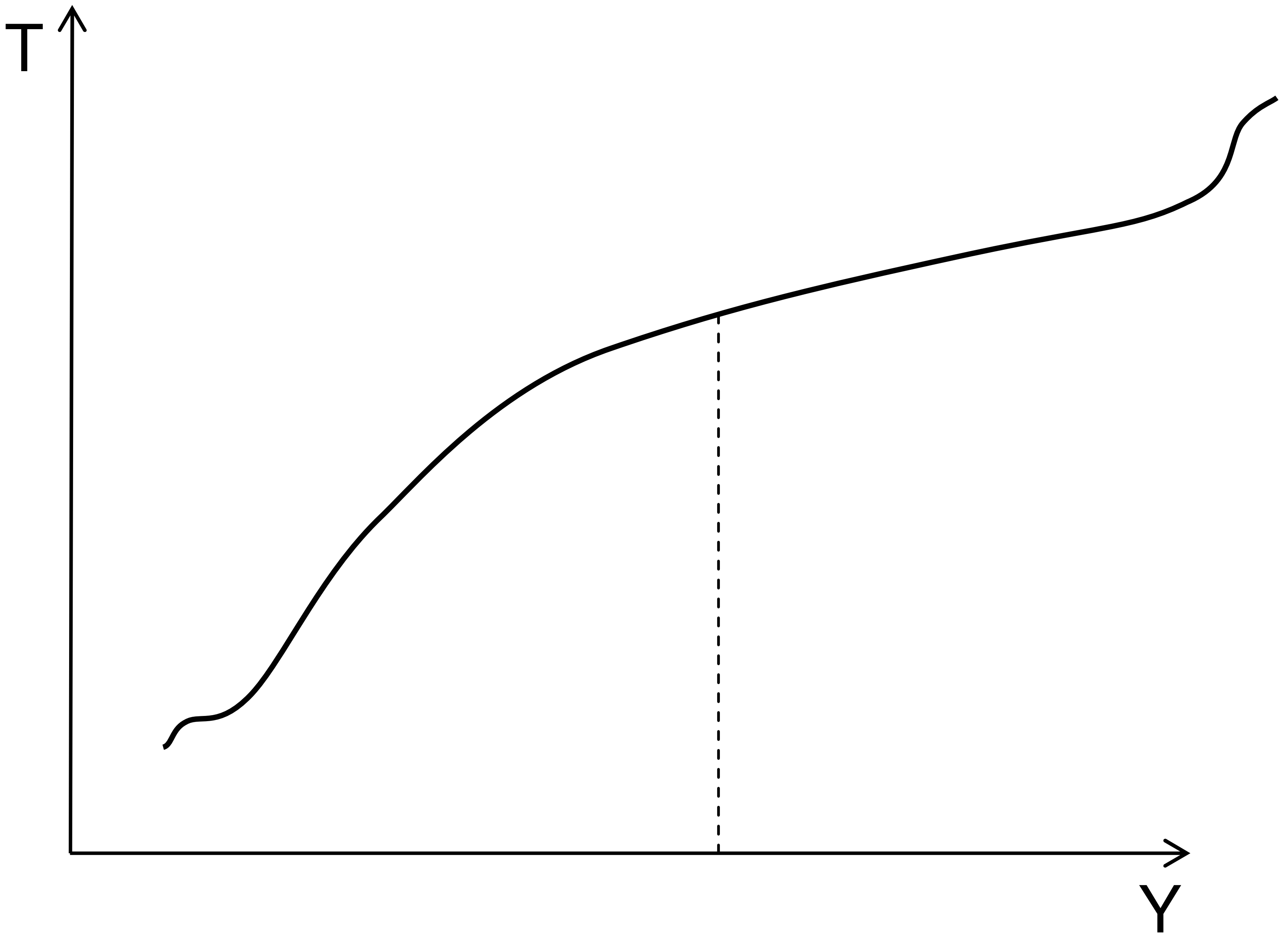
$$\hat{T}(Y) \equiv \begin{cases} T(Y(\hat{\theta})) - \varepsilon & Y = Y(\hat{\theta}) \\ T(Y) & Y \neq Y(\hat{\theta}) \end{cases}$$

Proposition. $\hat{T} \succ T$ ⟷

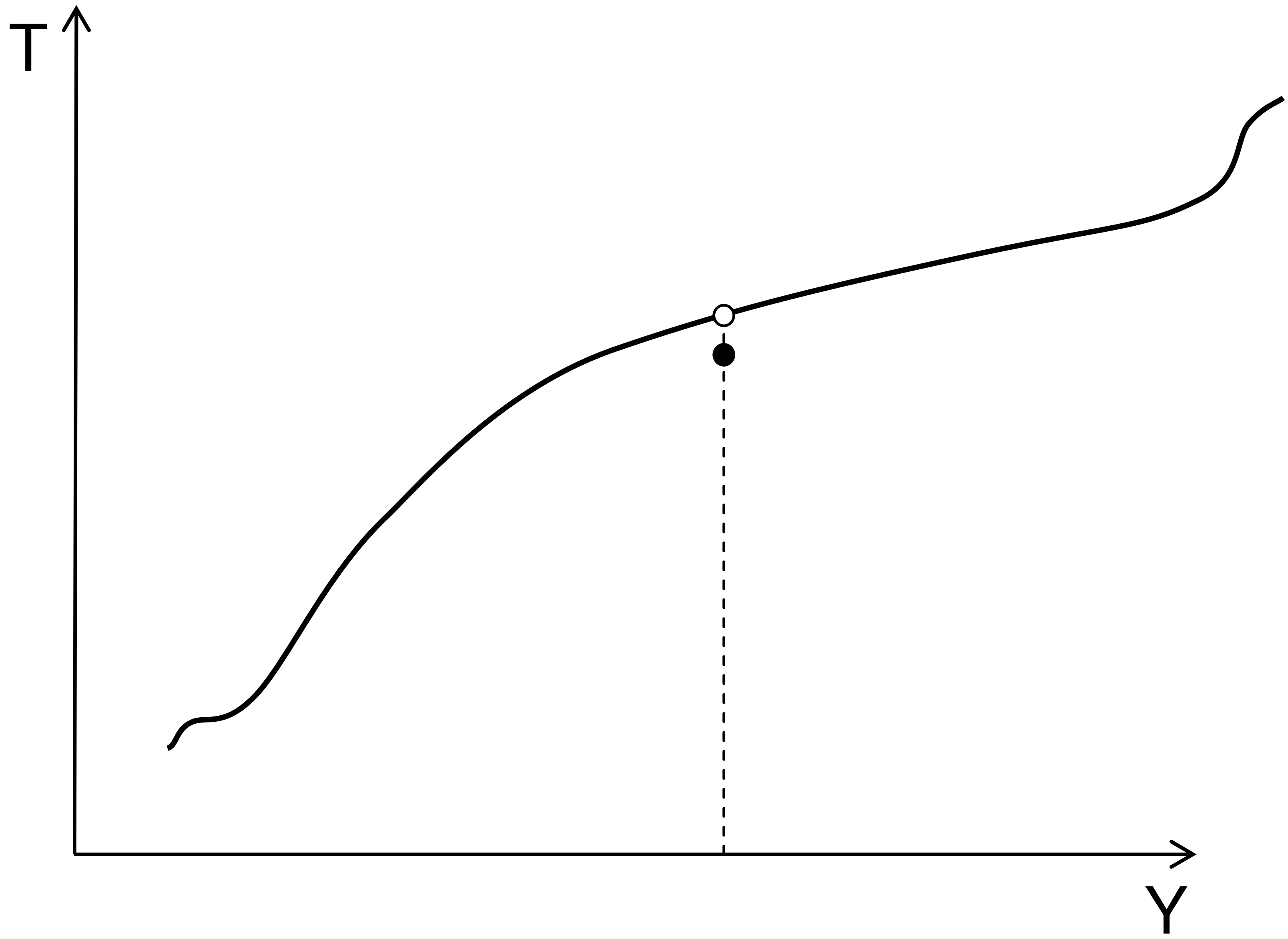
$$\tau(\theta)\left(\theta\frac{\tau'(\theta)}{\tau(\theta)} + 2\frac{d\log f(\theta)}{d\log\theta} - \frac{d\log h'(Y(\theta))}{d\log\theta}\right) \leq 3(1-\tau(\theta))$$

is violated at $\hat{\theta}$

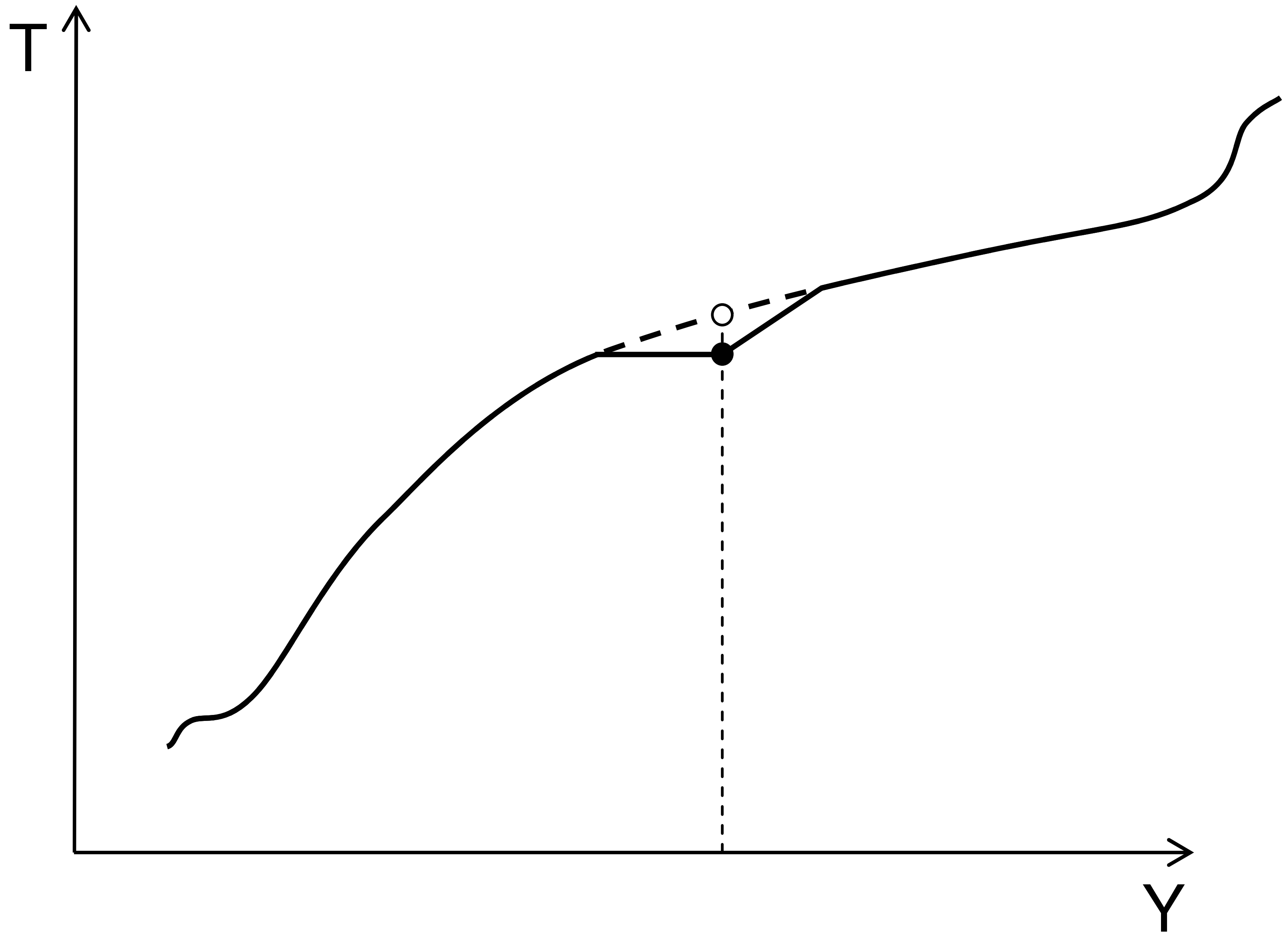
Simple Tax Reform



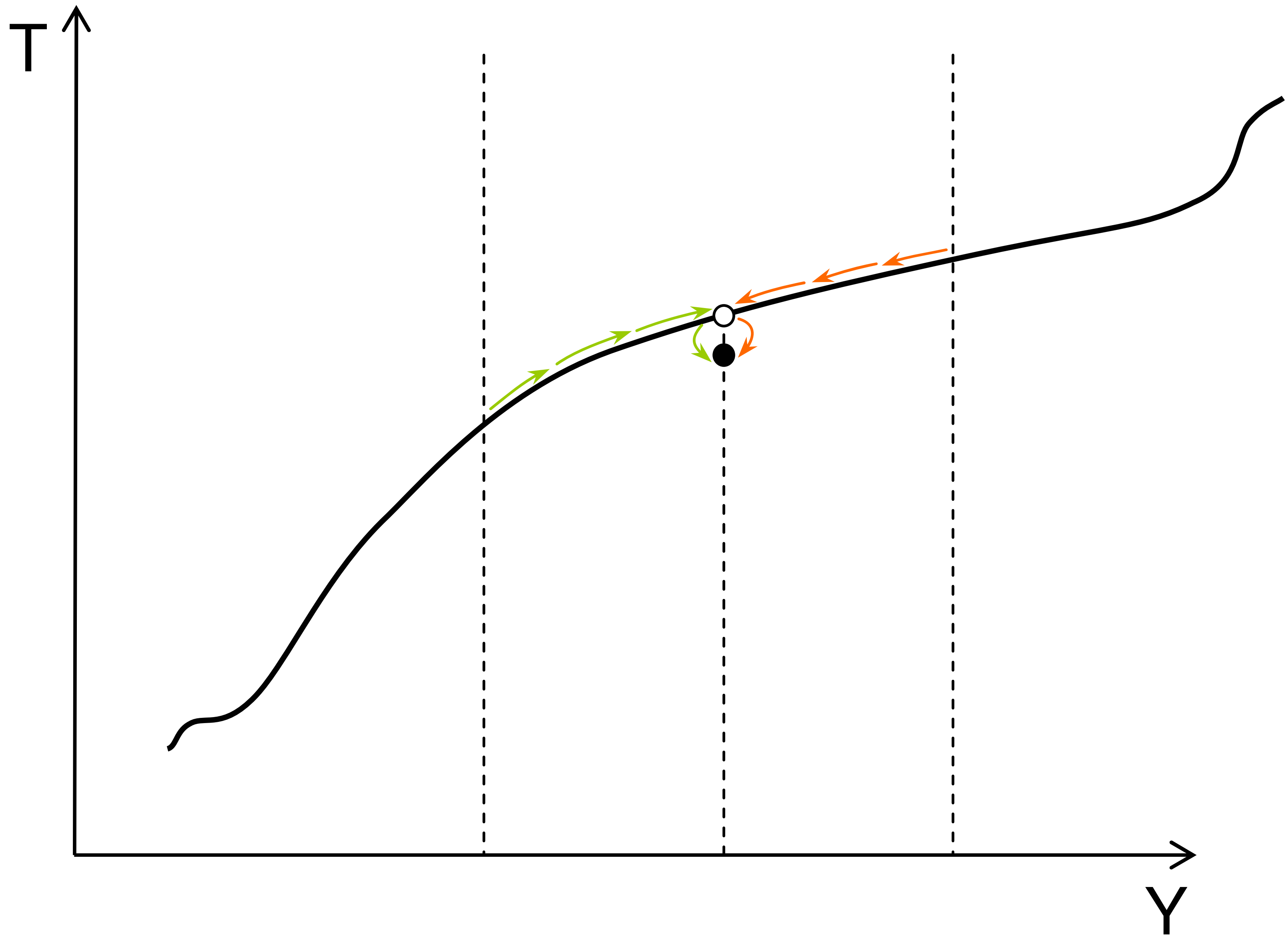
Simple Tax Reform



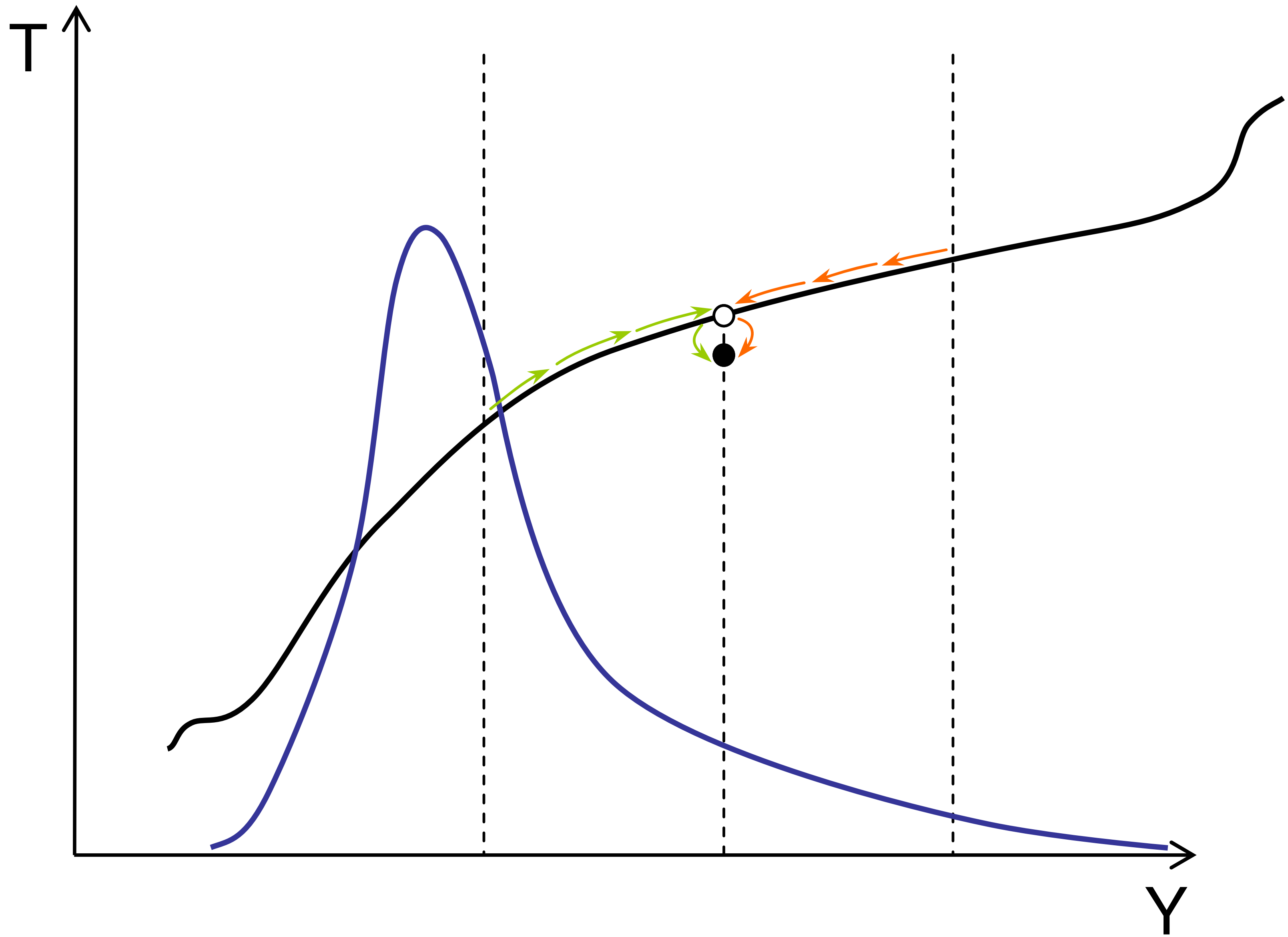
Simple Tax Reform



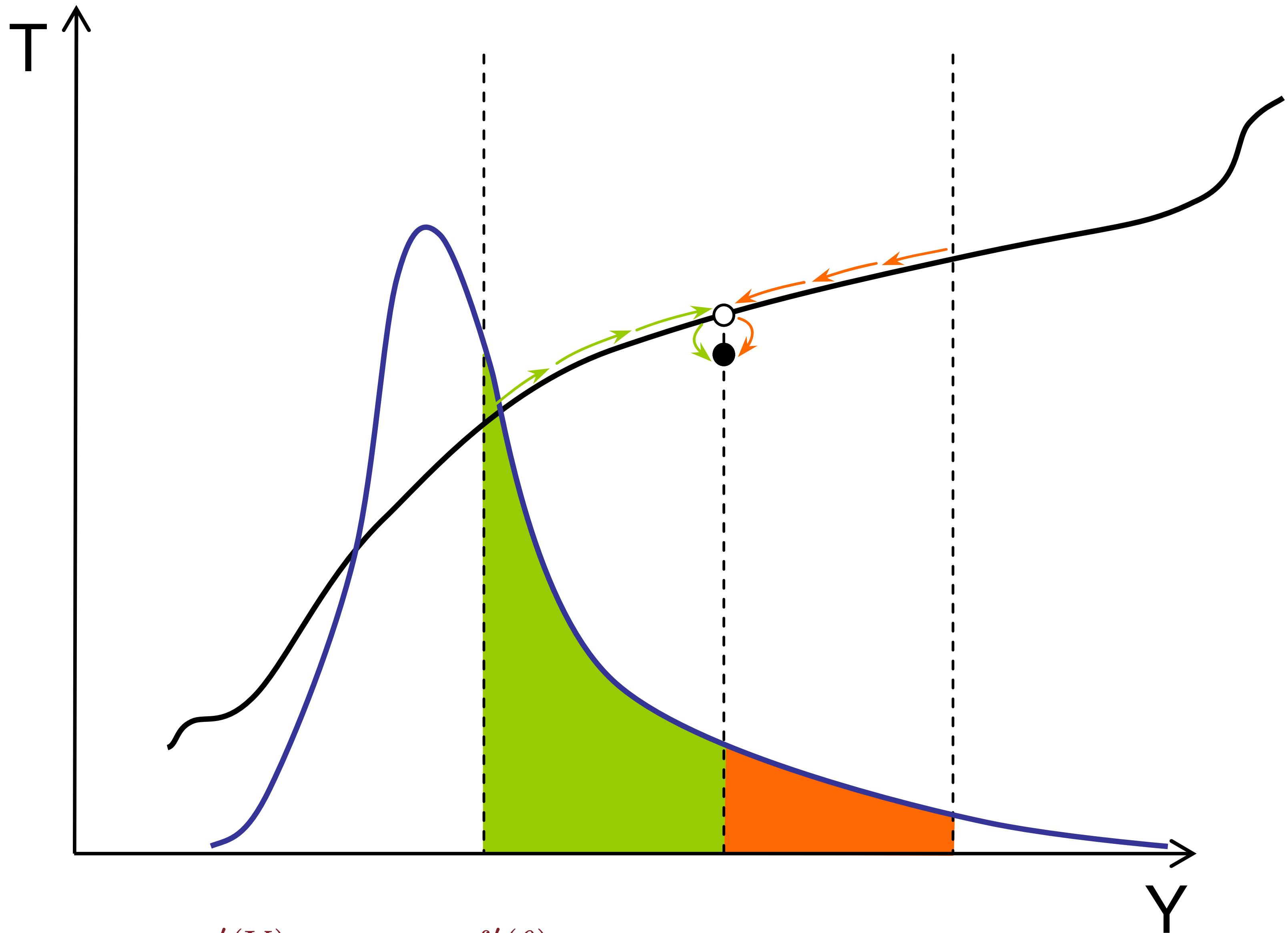
Simple Tax Reform



Simple Tax Reform



Simple Tax Reform



$\frac{g'(Y)}{g(Y)}$ small ($\frac{f'(\theta)}{f(\theta)}$ large) → inefficiency

Laffer

- lower taxes → increase revenue
- Pareto improvements ↔ "Laffer" effect

Proposition. $T_1(Y) \succ T_0(Y)$ → $T_1(Y) \leq T_0(Y)$

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

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

Proposition. For any $T(Y)$

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

- without empirical knowledge → anything goes
- need information on $f(\theta)$ to restrict $T(Y)$

Identification and Test

- observe $g(Y)$ identify (Saez, 2001)

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

$$\rightarrow f(\theta(Y)) = \frac{g(Y)}{\theta'(Y)}$$

Identification and Test

- observe $g(Y)$ identify (Saez, 2001)

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

$$\rightarrow f(\theta(Y)) = \frac{g(Y)}{\theta'(Y)}$$

- efficiency test...

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

... for tax schedule in place

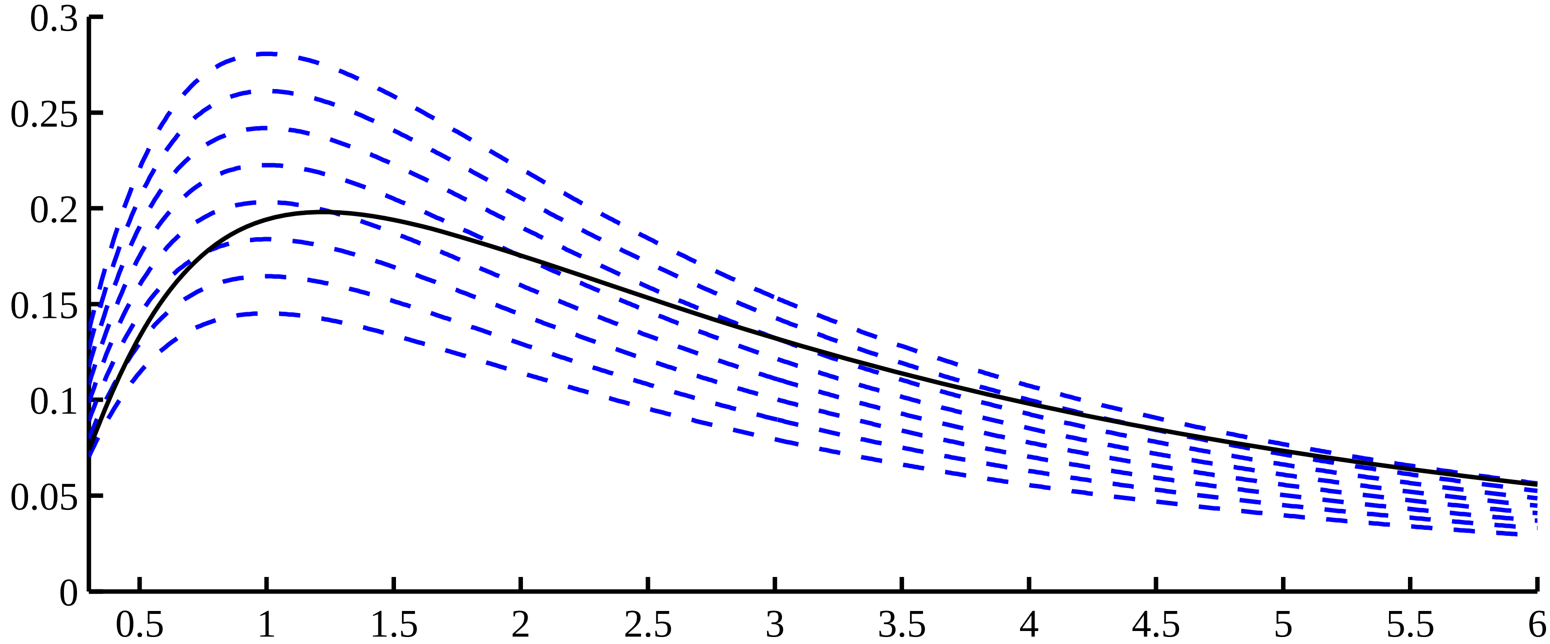
Graphical Test

- 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:

$$\frac{g(Y)}{\alpha(Y)} \qquad \text{nondecreasing}$$



Empirical Implementation

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

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)

- 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?
- $Y^i = \frac{1}{n} \sum Y_t^i$
- smooth density estimate
  assumed $T(Y) = .30 \times Y$

Output Density

- IRS's SOI Public Use Files for Individual tax returns
  - lifetime $g(Y)$?
  - lifetime $T(Y)$ schedule?
- $Y^i = \frac{1}{n} \sum Y_t^i$
- smooth density estimate assumed $T(Y) = .30 \times Y$

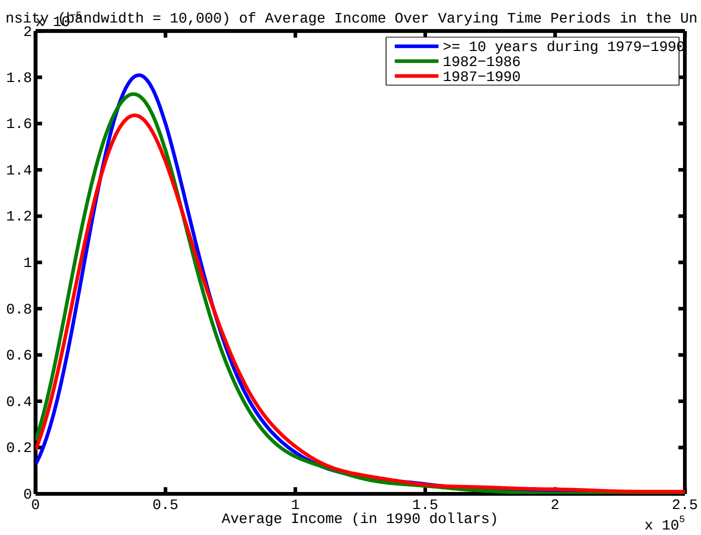


Figure 1: Density of income

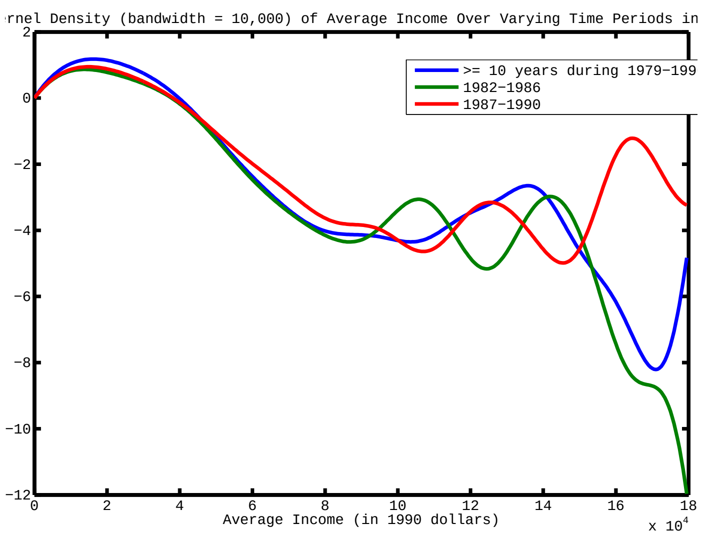
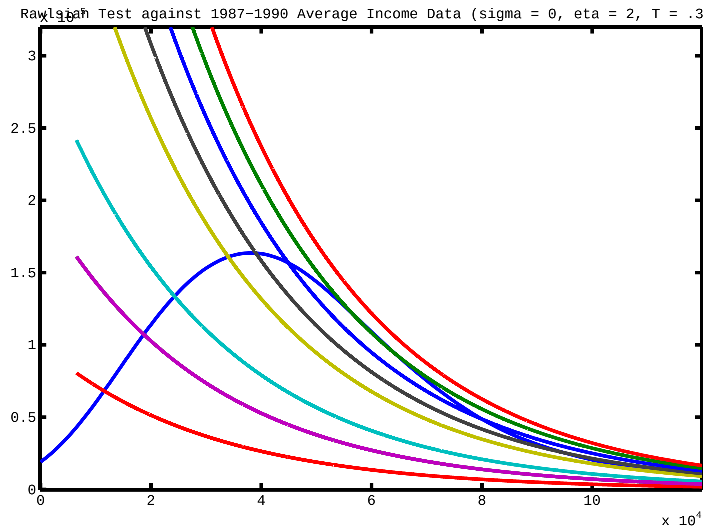
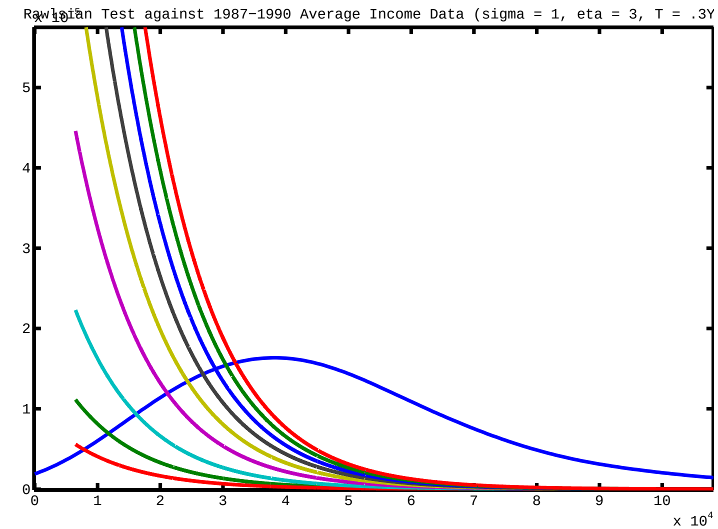


Figure 2: Implied elasticity

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 $T(Y) = .30 \times Y$





Quantifying Inefficiencies

- efficiency test → qualitative
- quantitative...

$$\Delta \equiv \int \left(\tilde{Y}^*(\theta) - \tilde{c}^*(\theta)\right) dF(\theta) - \int \left(Y(\theta) - c(\theta)\right) dF(\theta)$$

- 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

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

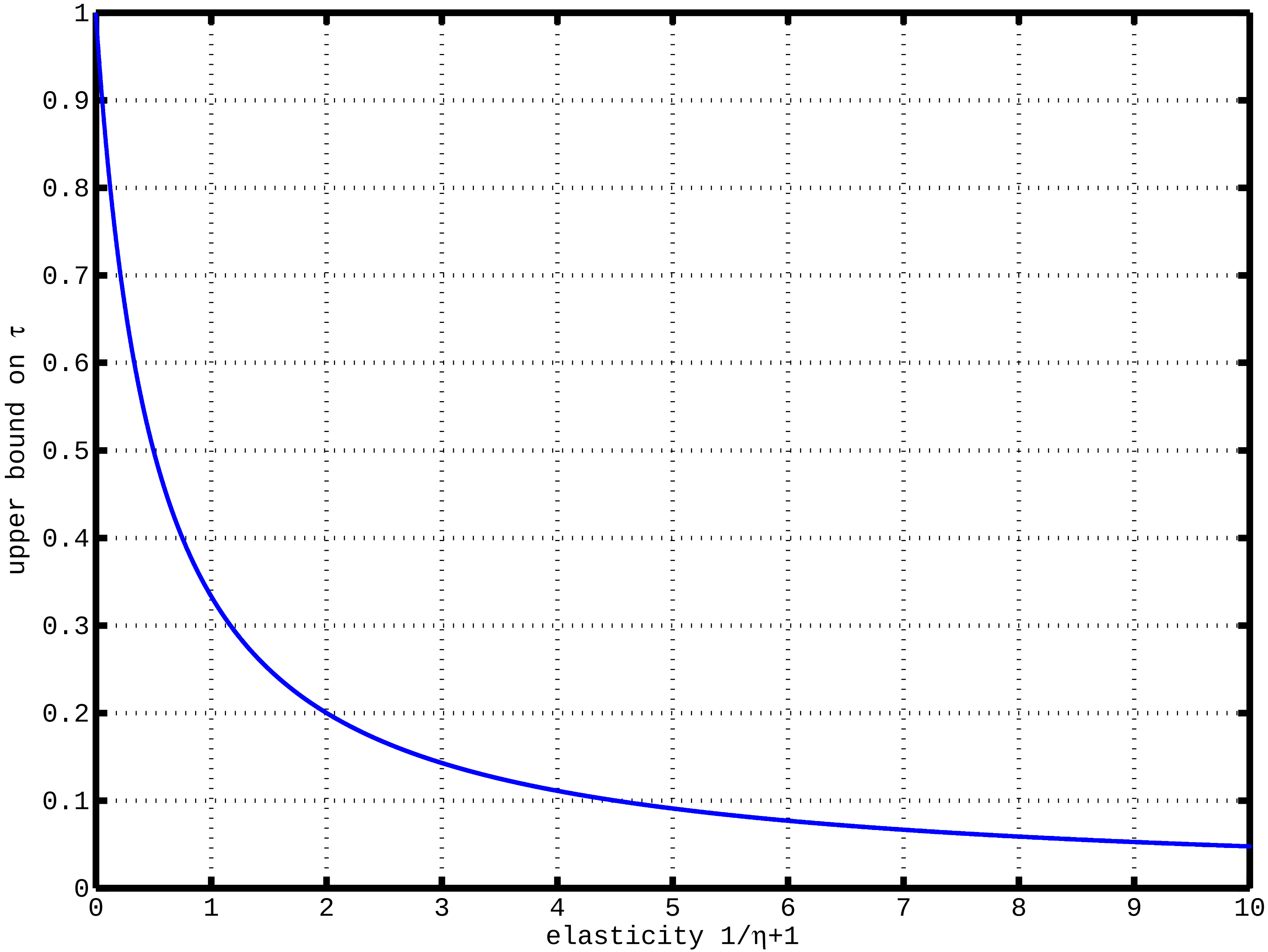
$$\bar{\tau} \equiv \lim_{\theta \to 0} \tau(\theta) = \lim_{Y \to \infty} T'(Y)$$

  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): $\varphi = 3$

Top Tax Rate



upper bound on τ
elasticity 1/η+1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7
8
9
10

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 $= 1, \ldots, N$

$$f^i(\theta) \qquad \text{and} \qquad U^i(c, Y, \theta)$$

- 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