Optimal Redistributive Taxation with both Extensive and Intensive Responses

Laurence JACQUET
THEMA - University of Cergy-Pontoise

Etienne LEHMANN
Université Panthéon Assas Paris 2 and CREST

Bruno VAN DER LINDEN
Université Catholique de Louvain and FNRS

Motivation

Economists rely intensively on constrained optimization.

In tax theory: max. "social welfare function" under constraints in order to determine optimal income tax profile.

An optimal income tax profile should:
(i) allow to maximize social welfare through income redistribution (i.e. taxes and transfers)
(ii) without inducing too many changes in individual behaviors (e.g. reducing their labor supply) that would affect the amount of collected taxes.

Motivation

Outline:

Presentation of a general optimal income tax model.

To discuss with you whether optimal transport could be useful to further generalize the tax model and its results.

Optimal labor income taxation: How best to solve the trade-off between equity and efficiency

- What is the best way to design labor income taxes given equity (social welfare) and efficiency concerns (taking into account tax-induced changes in behavior)?
- Seminal paper: Mirrlees (RES 1971)
- Information constraints (no way to make people reveal their individual characteristics at no cost)

  $\Rightarrow$ force to move to a world with inefficient/**distortionary** taxation, where tax rates depend on gross labor earnings, $T(Y)$ and not on individual characteristics.

Income taxation creates 2 types of distortion on the individual labor supply

- Responses along the **extensive** margin, i.e. people choose to work ($Y > 0$) or not ($Y = 0$)
This decision depends on **participation tax rates:** $\frac{T(Y)+b}{Y}$
where $b \geq 0$: welfare benefit for the non-employed.

- Responses along the **intensive** margin, i.e. workers choose "continuously" their labor earnings Y (their labor hours or effort)
This decision depends on **marginal tax rates:** $T'(Y)$

Previous results

- Model with intensive margin only and continuum of earnings (Mirrlees RES 1971): **Positive** marginal tax rates $T'(Y) \geq 0 \quad \forall Y$

- Model with extensive margin only and continuum of earnings (Diamond JPubE 1980) or with both margins and a finite number of earnings (using numerical simulations, Saez QJE 2002): **Negative** participation tax rates at the bottom $\frac{T(Y)+b}{Y} < 0$ may prevail, i.e. EITC

Redistribute more money to low-paid workers when they have a relatively large social welfare weight

Empirical evidence emphasizes the need to include both margins in the optimal tax literature

- Both labor supply margins empirically matter.

  See the large literature in microeconometrics, e.g., Gronau (JPE 1974), Heckman (Ecta 1974, 1979, AER 1993), Blundell and MaCurdy (1999 Handbook), Blundell, Bozio and Laroque (2011)

- However, so far, the literature does not say anything **analytically** about the signs of the **marginal** tax rates in a model with intensive and extensive margins simultaneously.

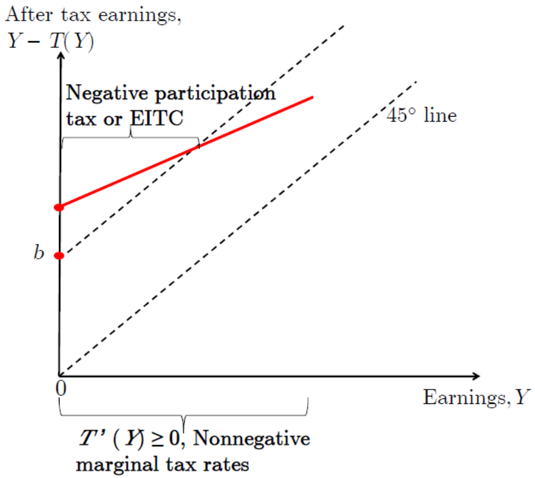
Our main contributions

A **new method to sign distortions** along the intensive margin in **any "random participation model"** (i.e. where the decision along the extensive margin depends on a random variable): Contribution to the self-selection literature à la Rochet and Stole (RES 2002).

Application in Income Taxation: When both intensive and extensive margins, under a fairly mild condition: $T'(.) > 0$ (Labor earnings are distorted downwards).

Our sufficient condition for positive marginal tax rates does not preclude negative participation tax rates (EITC)

Using US data, we show that negative participation tax rates (EITC) at the bottom and nonnegative marginal tax rates everywhere are optimal (under Benthamite preferences), i.e.



Agents differ along two dimensions and their preferences over consumption and effort (or earnings) are heterogeneous

- Fixed wage rates $\equiv$ skills: w

  on $[w_0, w_1]$ with $0 \leq w_0 < w_1 \leq +\infty$.

- Fixed disutilities of working (e.g., commuting, job search, reduced amount of time for home production) net of the stigma of being non-employed: χ

  on $(-\infty, \chi^{\max}]$ with $\chi^{\max} \leq +\infty$

- Joint density of types (χ, w): $k(\chi, w)$ continuous and positive over a connected support $(-\infty, \chi^{\max}] \times [w_0, w_1]$.
- The literature typically assumes that earnings are the product of the skill level times effort, $Y = w \times L$, we avoid this unnecessary restriction on the technology.

Preferences over consumption and effort (or earnings) are heterogeneous

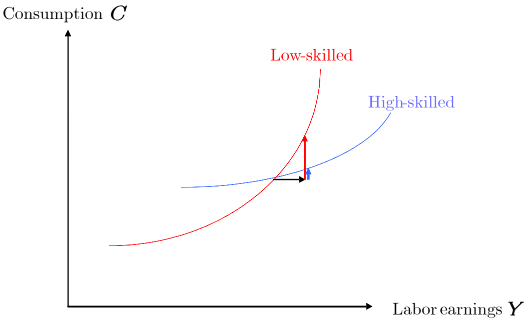
- Preferences over earnings, $Y > 0$, and consumption, $C = Y - T(Y)$:

$$\mathcal{U}(C, Y, w) - \mathbb{I}_{Y>0} \cdot \chi$$

with U twice-continuously differentiable and concave w.r.to (C, Y), $\mathcal{U}'_C > 0 > \mathcal{U}'_Y$ and $\mathcal{U}'_w > 0$ (i.e. a more skilled employed individual can get a given level of earnings by supplying a lower level of effort).

This utility allows **preferences over consumption and earnings to vary with skill** w.

- Strict single crossing (Spence-Mirrlees) condition $: w \mapsto \underbrace{-\frac{\mathcal{U}'_Y}{\mathcal{U}'_c}(C, Y, w)}_{\mathrm{MRS}_{CY}}$ is decreasing:



Intuition? Starting from any positive level of consumption C and earnings Y, to accept a unit rise in Y, more skilled workers need to be compensated with a smaller increase in C.

Individuals' choices

- Intensive choice: Employed workers choose different (positive) earnings Y because of skill heterogeneity:

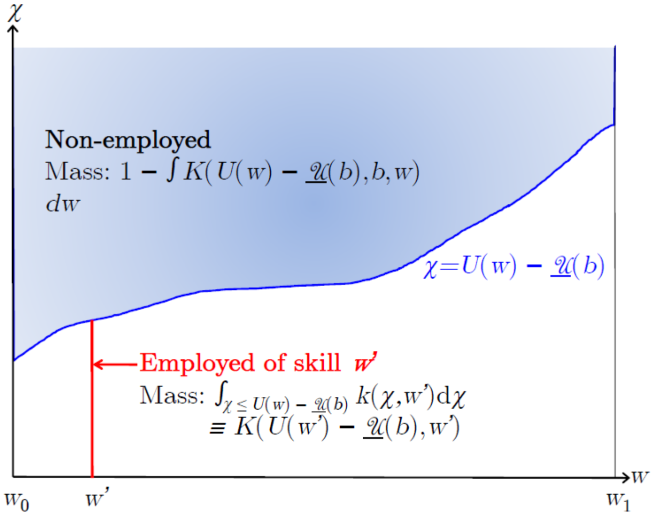
$$U(w) \equiv \max_{Y} \quad \mathcal{U}(Y - T(Y), Y, w)$$

- Extensive choice: Individuals of the same w take different participation decisions because they have different disutilities of participation χ:

  An individual (w, χ) chooses to work iff:

$$\begin{aligned} U(w) - \chi \quad &\geq \quad \mathcal{U}(b, 0, w) \equiv \underline{\mathcal{U}}(b) \\ \text{where } b \quad &: \quad \text{(endogenous) welfare benefit} \end{aligned}$$

At each skill level, we then have employed and non-employed people



Our Social Preferences generalize Bergson-Samuelson SWF

Sum a transformation $G\left(U, w, \chi\right)$ of individuals' utility U over all individuals:

$$\begin{aligned}\Omega &= \int_{w_0}^{w_1} \left\{ \int_{\chi \leq U(w) - \underline{\mathcal{U}}(b)} G\left(U\left(w\right) - \chi, w, \chi\right) \cdot k\left(\chi, w\right) d\chi \right. \\ &\quad \left. + \int_{\chi \geq U(w) - \underline{\mathcal{U}}(b)} G\left(\underline{\mathcal{U}}\left(b\right), w, \chi\right) \cdot k\left(\chi, w\right) d\chi \right\} dw\end{aligned}$$

$G'_U > 0$ and either $G''_{UU} < 0$ or $G''_{Uw} < 0$ or both.

New: G depends on utility levels as with Bergson-Samuelson SWF but also on the (w, χ)-type.
Relevant when the gvt wants to **compensate only for some characteristics (e.g. skill levels).**

Government's optimization problem

Find the optimal tax schedule T and welfare benefit b to maximize the previous objective function s.to:
(i) the government's budget constraint:

$$b = \int_{w_0}^{w_1} \left(Y\left(w\right) - C\left(w\right) + b \right) K\left(U\left(w\right) - \underline{\mathcal{U}}\left(b\right), w \right) \cdot dw - E$$

where E: exogenous amount of public expenditures.

(ii) the incentive-compatibility constraints: $\forall \left(w, x\right) \in \left[w_0, w_1\right]^2$
$U\left(w\right) = \mathcal{U}\left(C\left(w\right), Y\left(w\right), w\right) \geq \mathcal{U}\left(C\left(x\right), Y\left(x\right), w\right)$

These constraints impose that workers of skill w prefer the bundle $\left(C\left(w\right), Y\left(w\right)\right)$ designed for them rather than the bundle $\left(C\left(x\right), Y\left(x\right)\right)$ designed for workers of any other skill level x.

(ii) the incentive-compatibility constraints: $\forall (w, x) \in [w_0, w_1]^2$
$U(w) = \mathcal{U}(C(w), Y(w), w) \geq \mathcal{U}(C(x), Y(x), w)$

$\Leftrightarrow$
Y non-decreasing in w and $U'(w) = \mathcal{U}'_w(C(w), Y(w), w) > 0$.
(Mirrlees 1971)

$\Rightarrow$ **Optimal control** techniques can be used.

Marginal tax rates at the bottom and at the top

From FOC's:

- If skill distribution bounded ($w_1 < \infty$) and no bunching at the top: $T'(Y(w_1)) = 0$
- If $w_0 > 0$ and no bunching at the bottom: $T'(Y(w_0)) = 0$

i.e. same results as in the model with intensive margin only, Sadka (RES 1976), Seade (JPubE 1977).

Marginal tax rates for interior skill levels

A sufficient condition for **positive marginal tax rates** ($\forall w \in (w_0, w_1)$) thanks to a new method to sign distortions along the intensive margin.

Our method to sign distortions (valid in any model of monopoly screening with random participation):

- Characterize the optimum when the government observes the skills of workers but neither the skills of the non-employed nor the χ of anyone ("first-and-a-half-best"):

$$T\left(Y\left(w\right)\right)+b=\frac{1-g\left(w\right)}{\kappa\left(w\right)},$$

where $g\left(w\right)$: average mg social weight of w-workers,
$\kappa\left(w\right)$: participation response,
$\kappa\left(w\right)\equiv\frac{k}{K}\left(U\left(w\right)-\underline{\mathcal{U}}\left(b\right)\right)\cdot\mathcal{U}'\left(C\left(w\right),Y\left(w\right),w\right)$.

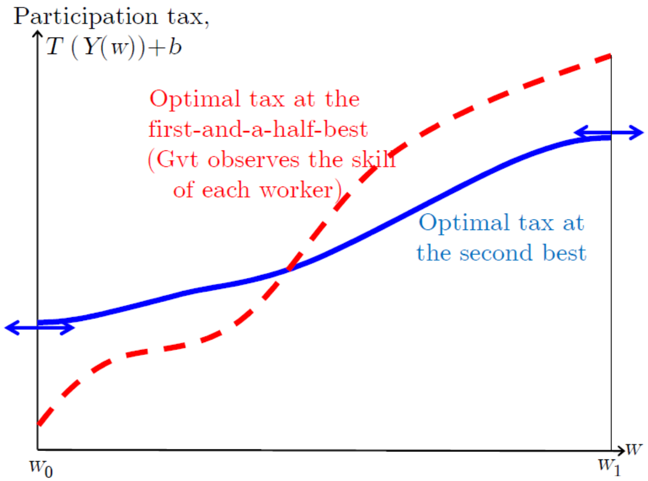
- Find a property on the first-and-a-half best optimum that allows to sign distortions along the intensive margin:

Property: $w\mapsto\frac{1-g(w)}{\kappa(w)}$ **admits a positive derivative everywhere** guarantees $T'(Y\left(w\right))>0$ in the first-and-a-half-best.

- We **show** that this property is still valid when the government does not observe the skills and the χ (second-best).

A sufficient condition for nonnegative mg tax rates:

Along the second-best allocation, $w \mapsto \frac{1-g(w)}{\kappa(w)}$ **admits everywhere a positive derivative is sufficient** for $T'(Y(w)) > 0$



Examples on primitives where marginal tax rates are positive and participation tax rates are nonnegative (NIT)

Assume:

- Additive separable utility function

$$u\left(C\right) - v\left(Y, w\right) - \mathbb{I}_{Y>0} \cdot \chi$$

- If $w \mapsto k\left(\chi, w\right) / K\left(\chi, w\right)$ admits everywhere a negative partial derivative in χ and a non-positive partial derivative in w

- **Either** Maximin

  or Benthamite social preferences and $g\left(w_0\right) \leq 1$

  $\Rightarrow T'(Y\left(w\right)) > 0$ and NIT.

Applications in other frameworks with random participation: Regulation of a monopoly and non-linear pricing

	Informed Party (Agents)	Uninformed Party (Principal)	Unobservable characteristics	Intensive and extensive margins
Labor Income Taxation	Workers	Tax authority	• Disutilities when working • Skills	• Whether to work • How much
Regulation of a Monopoly	Monopoly	Regulator	• Fixed cost • Marginal cost	• Whether to produce • How much
Non-linear Pricing	Consumers	Firm	• Outside options • Tastes	• Whether to buy • How much

With the same assumptions on the primitives: downward distortions in other adverse selection problems

- **Regulatory monopoly**: Regulated price > Marginal cost (Production is distorted downwards)
- **Nonlinear pricing**: Marginal price >Marginal cost of production (Quantity bought is distorted downwards)

We check the empirical relevance of our condition using US data

- Social preferences: Bentham and Maximin
- Individuals' preferences:

$$\mathcal{U}(C, Y, w) = \frac{\left(C - \left(\frac{Y}{w}\right)^{1+\frac{1}{\varepsilon}} + 1\right)^{1-\sigma}}{1-\sigma}$$

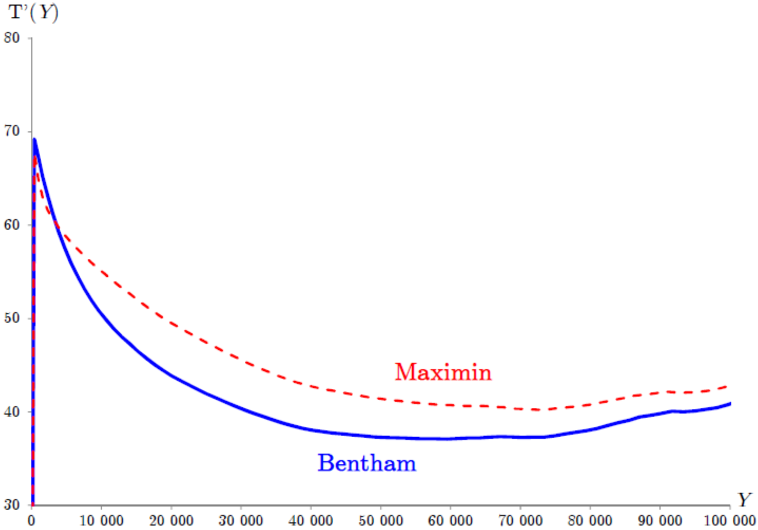
  No income effect along the intensive margin
- $\varepsilon = 0.25$ in the benchmark (In the US, $\varepsilon \in [0.12, 0.4]$, Saez et al. 2010)
- $\sigma = 0.8$ in the benchmark
- We use weakly earnings in 2007 CPS for singles (without kids)

- The skill distribution among employed workers is calibrated such that the actual $T(.)$ yields **empirical earnings**
- The skill density is smooth, using a **quadratic kernel** (bandwidth \$3822)
- The **top** (3.3%) of the distribution is approximated by a **Pareto distribution** with Pareto index $a = 2$ following Diamond (*AER* 1998) and Saez (*RES* 2001)
- Remark: $w_0 = \$0.1$ (that corresponds to an annual $Y(w_0) < \$1$)
- CDF of χ, conditional on skill level is logistic:

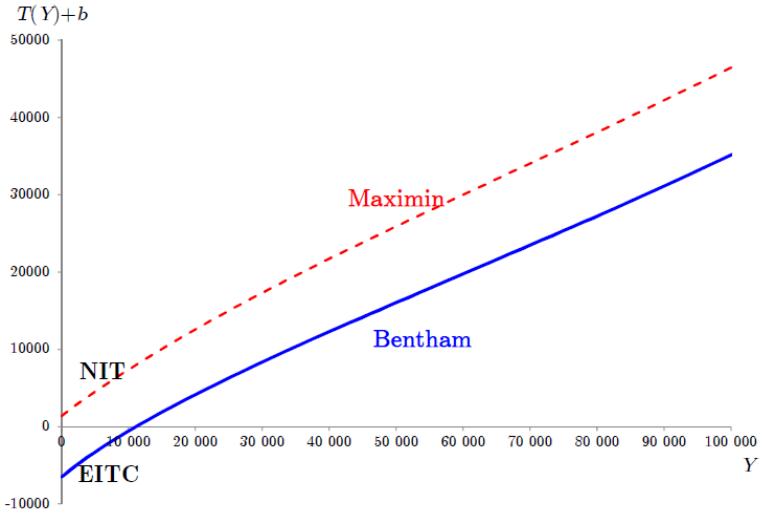
$$K(\chi, w) = \frac{\exp\left(-\alpha(w) + \beta(w)\ \chi\right)}{1 + \exp\left(-\alpha(w) + \beta(w)\ \chi\right)}$$

  Parameters $\alpha(w)$ and $\beta(w)$ are calibrated to obtain **skill-specific** employment rates and **skill-specific** elasticities of employment rates w.r.to $C(w) - b$ that are empirically relevant.

Optimal U-shape profiles for the marginal tax rates

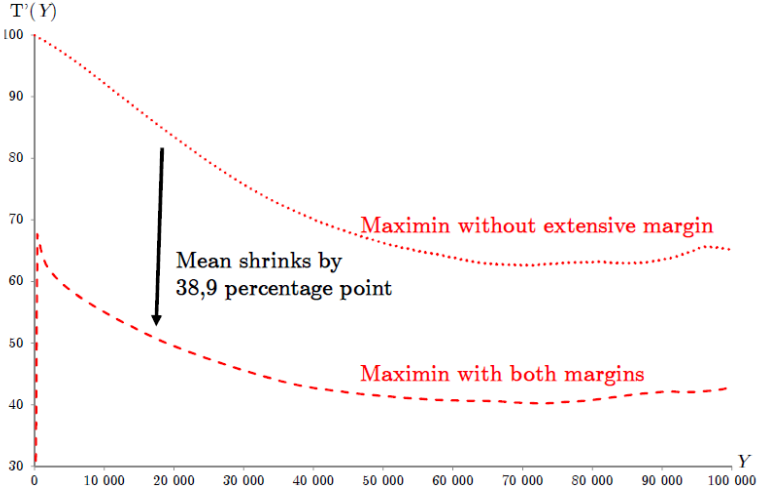


NIT under Maximin and EITC under Bentham



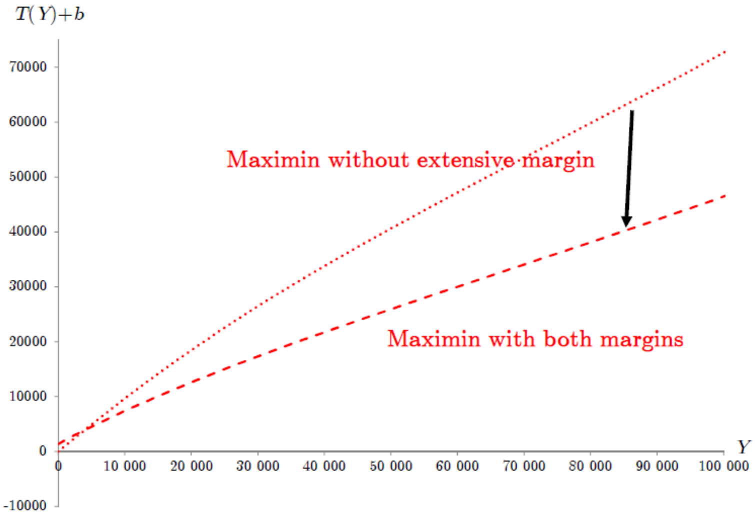
Introducing the extensive margin drastically reduces the optimal marginal tax rates

e.g., under Maximin:

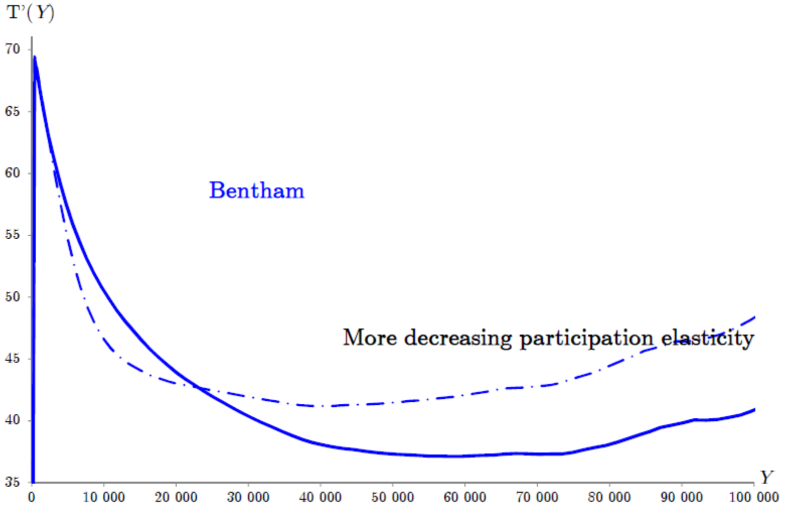


Introducing the extensive margin drastically modifies the optimal participation taxes

e.g., under Maximin:



A more decreasing participation elasticity shifts the marginal tax rates upwards



Sensitivity exercises

All our sensitivity exercises show:

- Marginal tax rates always positive (and nil at the two extremes)
- **Our sufficient condition always holds**
- Marginal tax rate are always U-shaped

Synthesis

- Quite complete model of optimal taxation: 2 margins, continuum of skills and general individual and social preferences
- A sufficient condition for nonnegative marginal tax rates + Examples
- New method to sign distortions along the intensive margin that can largely be used in other screening models with random participation
- Robustness of our condition + importance of extensive margin on optimal tax schedule by implementing the model with U.S. data
- Nonnegative marginal tax rates may coexist with negative participation tax rates at the bottom
- Under Bentham: EITC vs under Maximin: NIT

Could optimal transport be helpful to relax some assumptions used in tax theory?

The intensive labor supply decision (Y) of individuals that have chosen to work depend only on their skill and not on their net disutility of participation χ.

The strict Spence-Mirrlees single crossing condition.

Using Optimal Transport, is it still possible to characterize marginal tax rates in a more general model?