Public Finance

Pierre Boyer

École polytechnique - CREST

Fall 2025

Outline of the class

Overview of Optimal Taxation

- From an efficiency perspective, would finance government purely through lump-sum taxation
- With redistributional concerns, would ideally levy individual-specific lump sum taxes
  - Tax higher-ability individuals a larger lump sum
- Problem: cannot observe individuals' types
  - Therefore must tax economic outcomes such as income or consumption, which leads to distortions

Ramsey vs. Mirrleesian Approaches

- Two approaches to optimal taxation
  1. Ramsey: restrict attention linear tax systems: $t \cdot x$
  2. Mirrleesian: Non-linear tax systems, with no restrictions on $T(x)$
- Ramsey approach: rule out possibility of lump sum taxes by assumption and consider linear taxes
- Mirrleesian approach: permit lump sum taxes, but model their costs in a model with heterogeneity in agents' skills

Central Results in Tax Theory covered in Part II

1. Ramsey (1927): inverse elasticity rule
2. Chamley (1985), Judd (1986): no capital taxation in infinite horizon Ramsey models
3. Diamond and Mirrlees (1971): production efficiency
4. Atkinson and Stiglitz (1976): no consumption taxation with optimal *non-linear* income taxation

Optimal Income Taxation

1. Optimal Static Income Taxation: Mirrlees (1971)
2. Empirical Implementation of Mirrlees Model: Diamond (1998), Saez (2001)
3. Income and Commodity Taxation: Atkinson and Stiglitz (1976) [Part II]
4. Optimal Transfer Programs: Saez (2002) [Part II]

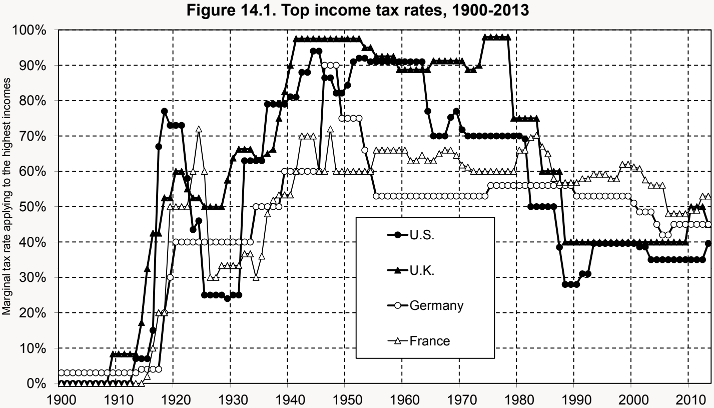
Income tax brackets

Tax $T(z)$ is piecewise linear and continuous function of taxable income z with constant marginal tax rates (MTR) $T'(z)$ by brackets

- Introduced in 1913, Federal US income tax in 2024 has 7 brackets with MTR 10%, 12%, 22%, 24%, 32%, 35%, 37% (top bracket for z above \$609K single)

- Introduced in 1914, France Impôt sur le Revenu in 2022 has 4 brackets with MTR 11%, 30%, 41%, 45% (top bracket for z above 180K euros, exemption for z below 11K euros)
- History of introduction of income tax in France: Les Batailles de l'impôt by Nicolas Delalande (Seuil, 2011) & Piketty (2001, 2018; Les Hauts revenus en France au 20e siècle) & Boyer (PUF, 2024)

Tax rates change frequently over time. Statutory top MTRs have declined drastically since 1960s in many OECD countries.



The top marginal tax rate of the income tax (applying to the highest incomes) in the U.S. dropped from 70% in 1980 to 28% in 1988. Sources and series: see piketty.pse.ens.fr/capital21c.

Main means-tested transfer programs

1. Traditional transfers: managed by welfare agencies, paid on monthly basis, high stigma and take-up costs

⇒ low take-up rates

Main programs

- US: Medicaid (health insurance for low incomes), SNAP (former food stamps), public housing, TANF (traditional welfare), SSI (aged+disabled)
- France: sécurité sociale, public housing (HLM), RSA

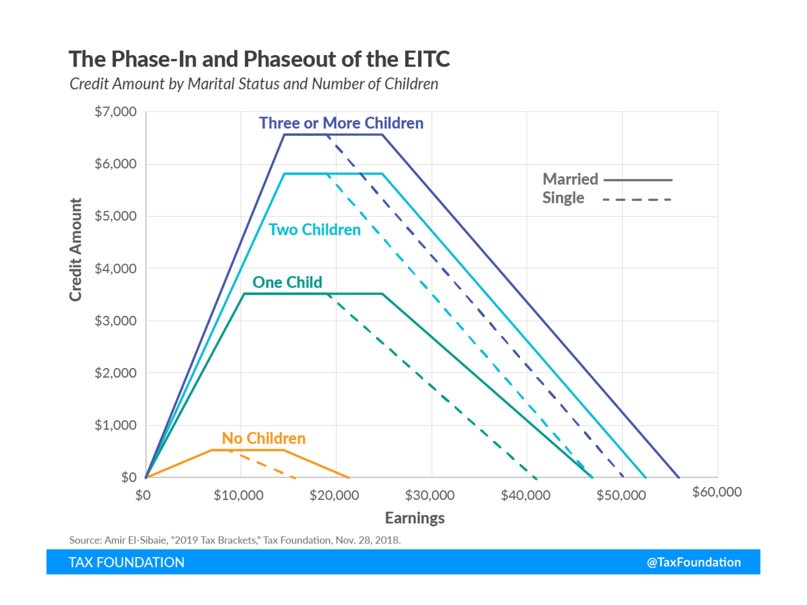
Main means-tested transfer programs

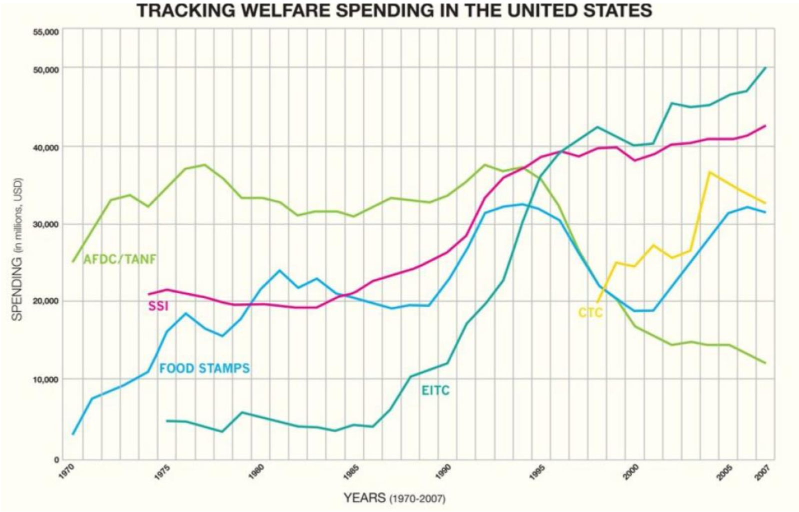
2. Refundable income tax credits: managed by tax administration, paid as an annual lumpsum in year $t + 1$, low stigma and take-up cost

$\Rightarrow$ high take-up rates

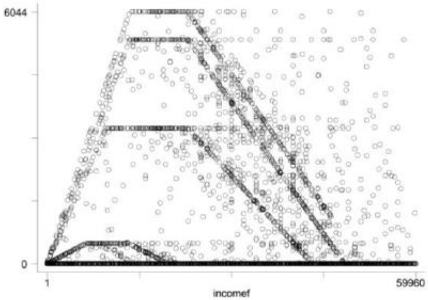
Main programs:

- US: EITC and Child Tax Credit (large expansion since the 1990s) for low income working families with children
- In France it was prime pour l'emploi: now prime d'activité (merger of prime pour l'emploi+ RSA since 2016) managed by welfare agency (CAF).
- Germany: proposal ifo & ZEW team (Peichl et al., 2023).


The Phase-In and Phaseout of the EITC
Credit Amount by Marital Status and Number of Children
$7,000
$6,000
$5,000
$4,000
$3,000
$2,000
$1,000
$0
Credit Amount
Three or More Children
Two Children
One Child
No Children
Married
Single
$0
$10,000
$20,000
$30,000
$40,000
$50,000
$60,000
Earnings
Source: Amir El-Sibaie, "2019 Tax Brackets," Tax Foundation, Nov. 28, 2018.
TAX FOUNDATION
@TaxFoundation

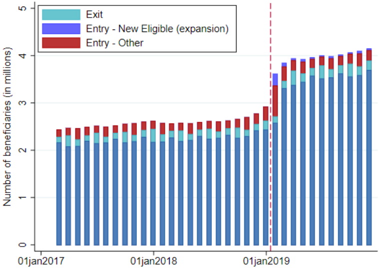


TRACKING WELFARE SPENDING IN THE UNITED STATES
55,000
50,000
40,000
30,000
20,000
10,000
0
SPENDING (in millions, USD)
AFDC/TANF
SSI
FOOD STAMPS
EITC
CTC
1970
1975
1980
1985
1990
1995
2000
2005
2007
YEARS (1970-2007)



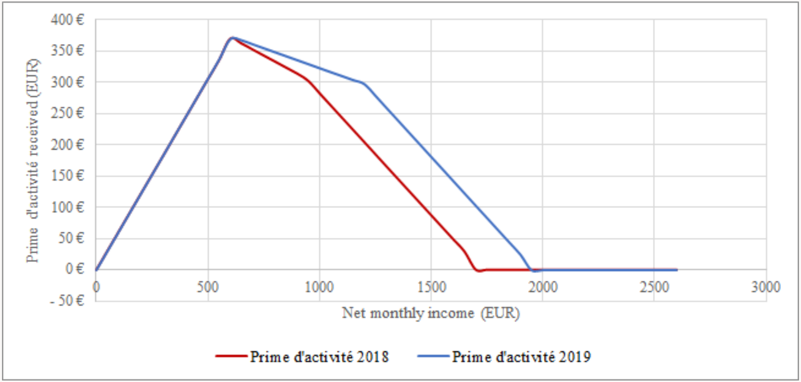
Perceived EITC for married individuals in 2014

Leroy (2023)



Source: Social records (ALLSTAT) and author's own computations.
Note: New eligible due to the benefit's eligibility expansion are identified using micro-simulation techniques. Other entrants can reflect a change in eligibility (unrelated to the benefit's expansion) or a change in take-up behaviors.

Note IPP: Boyer et al. (2024)



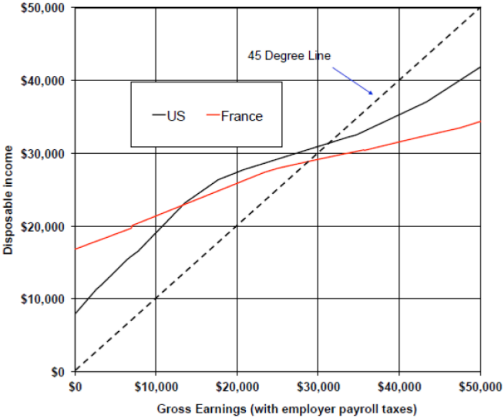
Key concepts for taxes/transfers

1. Transfer benefit with zero earnings $-T(0)$ (sometimes called demogrant or lumpsum grant)
2. Marginal tax rate (or phasing-out rate) $T'(z)$: individual keeps $1 - T'(z)$ for an additional \$1 of earnings (intensive labor supply response)
3. Participation tax rate $\tau_p = [T(z) - T(0)]/z$: individual keeps fraction $1 - \tau_p$ of earnings when moving from zero earnings to earnings z (extensive labor supply response):

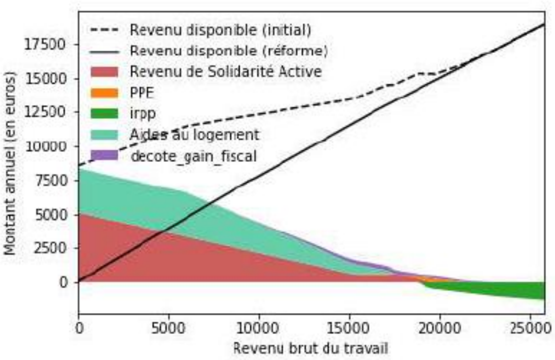
$$z - T(z) = -T(0) + z - [T(z) - T(0)] = -T(0) + z \cdot (1 - \tau_p)$$

4. Break-even earnings point z^*: point at which $T(z^*) = 0$

- The intensive margin: Number of hours of work (or intensity of work) of participating workers
- The extensive margin: Participation decision [Part II]



Source: Piketty, Thomas, and Emmanuel Saez (2012)


Revenu disponible (initial)
Revenu disponible (réforme)
Revenu de Solidarité Active
PPE
irpp
Aides au logement
decote_gain_fiscal
Montant annuel (en euros)
17500
15000
12500
10000
7500
5000
2500
0
0
5000
10000
15000
20000
25000
Revenu brut du travail

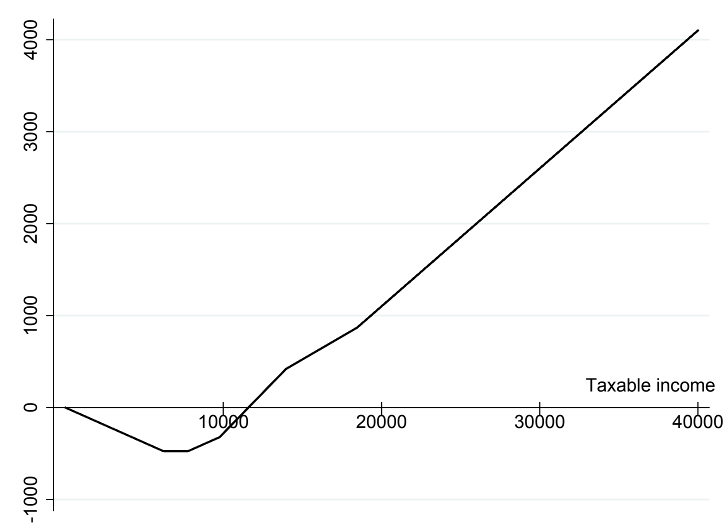


Figure: Income tax schedules for singles without dependants from micro-simulation models for the US in 2012 (Bierbrauer and Boyer, 2018)

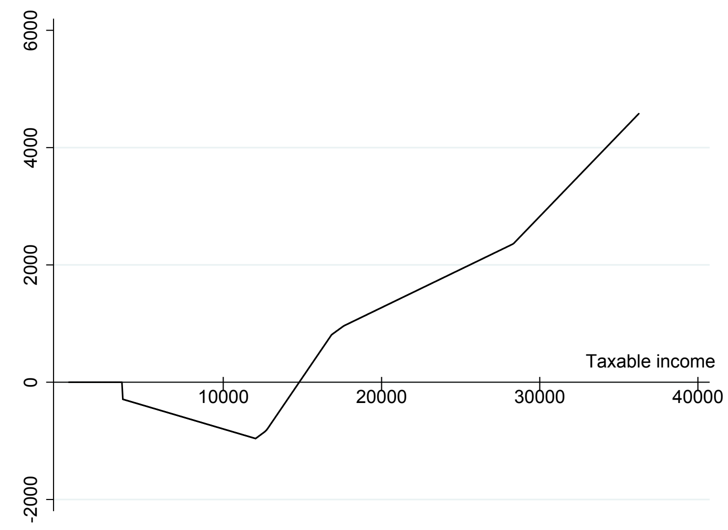
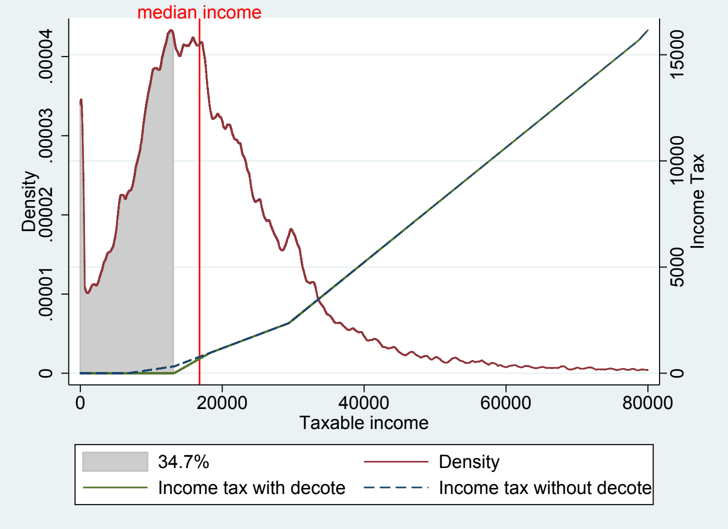


Figure: Income tax schedules for singles without dependants from micro-simulation models for France in 2012 (Bierbrauer and Boyer, 2018)


median income
Density
0
.00001
.00002
.00003
.00004
0
5000
10000
15000
Income Tax
0
20000
40000
60000
80000
Taxable income
34.7%
Density
Income tax with decote
Income tax without decote

Optimal taxation: simple model with no behavioral responses

Utility $u(c)$ strictly increasing and concave

Same for everybody where c is after tax income.

Income is z and is fixed for each individual, $c = z - T(z)$ where $T(z)$ is tax on z. z has density distribution $h(z)$

Government maximizes **Utilitarian** objective:

$$\int_0^\infty u(z - T(z))h(z)dz$$

subject to **budget constraint** $\int T(z)h(z)dz \geq E$ (multiplier λ)

Simple model with no behavioral responses

Form lagrangian: $L = [u(z - T(z)) + \lambda T(z)]h(z)$

First order condition (FOC) in $T(z)$:

$$0 = \frac{\partial L}{\partial T(z)} = [-u'(z - T(z)) + \lambda]h(z) \Rightarrow u'(z - T(z)) = \lambda$$

$\Rightarrow z - T(z) =$ constant for all z.

$\Rightarrow c = \bar{z} - E$ where $\bar{z} = \int zh(z)dz$ average income.

100% marginal tax rate. Perfect equalization of after-tax income.

Utilitarianism with decreasing marginal utility leads to perfect egalitarianism (Edgeworth, 1897)

Issues with simple model

1. **No behavioral responses:** Obvious missing piece: 100% redistribution would destroy incentives to work and thus the assumption that z is exogenous is unrealistic

$\Rightarrow$ Optimal income tax theory incorporates behavioral responses (Mirrlees, 1971): **equity-efficiency trade-off**

2. **Issue with Utilitarianism:** Even absent behavioral responses, many people would object to 100% redistribution (perceived as confiscatory)

⇒ Citizens' views on fairness impose **bounds** on redistribution gov't can do (political economy / public choice theory)

⇒ How preferences for redistribution are shaped? Alesina, Giuliano, Bisin, and Benhabib (2011, Handbook of Social Economics).

- Discussion of alternatives to utilitarianism: in economics (Saez and Stantcheva - Weinzierl - Fleurbaey and Maniquet) but also in political philosophy (Sandel, 2010. Justice: What's the Right Thing to Do?).
- Now assumed generalized utilitarian planner ⇒ we come back to the issue in next set of slides.

Mirrlees model results

- Optimal income tax trades-off redistribution and efficiency (as lump sum taxes not feasible)

$\Rightarrow$ $T(.) < 0$ at bottom (transfer) and $T(.) > 0$ further up (tax) (full integration of taxes/transfers)

- Mirrlees formulas complex, only a couple fairly general results:
  1. $0 \leq T'(.) \leq 1$, $T'(.) \geq 0$ is non-trivial (rules out EITC)
  2. Marginal tax rate $T'(.)$ should be zero at the top (if skill/income distribution bounded)

Beyond Mirrlees

- Mirrlees (1971) had a huge impact on information economics: models with asymmetric information in contract theory
- Discrete 2-type version of Mirrlees model developed by Stiglitz (1982) with individual FOC replaced by Incentive Compatibility constraint (high-type individuals should not mimic low-type individuals)

- Till late 1990s, Mirrlees results not closely connected to empirical tax studies and little impact on tax policy recommendations

- Since late 1990s, Piketty (1997), Diamond (1998), Saez (2001) have connected Mirrlees model to practical tax policy / empirical tax studies

New approach summarized in Diamond and Saez (2011), Boadway (2012), and Piketty and Saez (2013, Handbook Chapter).

Mirrlees General Setup: Preferences I

There is a continuum of individuals of measure 1. The set of individuals is denoted by I, with generic entry i.

Individuals consume and contribute to the economy's output/ generate income. If individual i consumes c_i and has an output requirement of y_i, she realizes utility of

$$u(c_i, y_i, \omega_i) \ ,$$

where ω_i is a productivity parameter which belongs to a set $\Omega \subset \mathbb{R}_+$. We refer to ω_i also as individual i's type.

Mirrlees General Setup: Preferences II

The function u is increasing in its first argument and decreasing in its second argument: $u_1(\cdot) \geq 0$ and $u_2(\cdot) \leq 0$.

We assume that, *for any pair (c, y) the marginal rate of substitution between consumption and income is decreasing in the individual's type*. Formally, for all $(c, y) \in \mathbb{R}^2_+$, and for any pair $\omega, \omega' \in \Omega$, with $\omega' > \omega$

$$-\frac{u_2(c, y, \omega)}{u_1(c, y, \omega)} > -\frac{u_2(c, y, \omega')}{u_1(c, y, \omega')}$$

This property is referred to as the Spence-Mirrlees single crossing property.

Mirrlees General Setup: Preferences III

In a y-c-diagram, the individual's indifference curves are increasing. The Spence-Mirrlees single crossing property implies that higher types/ more productive individuals have indifference curves that are flatter: An individual that is more productive requires less compensation for a marginal increase of the output requirement.

The cross-section distribution of types is represented by a cumulative distribution function F with density f.

Remark

It is often assumed that there is a function $U : \mathbb{R}^2_+ \to \mathbb{R}$ so that

$$u(c, y, \omega) = U\left(c, \frac{y}{\omega}\right) \; .$$

The interpretation is as follows: Individuals value consumption and leisure. Therefore utility is increasing in consumption, the first argument of U, and decreasing in hours worked, the second argument of U. Moreover, an individuals' type is identified with his hourly wage. Thus, to generate an income of y, an individual with type ω needs to work for $\frac{y}{\omega}$ hours.

Admissible Allocations I

An allocation consists of a function $c : \Omega \to \mathbb{R}_+$ and a function $y : \Omega \to \mathbb{R}_+$, which specify, respectively, the individuals' consumption levels and output requirements as a function of their types.

Definition: Incentive compatibility of an allocation

An allocation $c : \Omega \to \mathbb{R}_+, y : \Omega \to \mathbb{R}_+$ is said to be incentive compatible if, for any pair $\omega, \omega' \in \Omega$,

$$u(c(\omega), y(\omega), \omega) \geq u(c(\omega'), y(\omega'), \omega) .$$

Admissible Allocations II

Interpretation of the requirement of incentive-compatibility:

- Suppose that individuals have private information on their type.
- A planner wishes to implement an allocation where consumption levels and output requirements are type-dependent.
- The planner therefore engages in a communication game: Each individual reports a type and gets the consumption level and the output requirement which have been specified for the reported type.

Admissible Allocations III

- The incentive compatibility constraints ensure that individuals are willing to reveal their types and therefore end up with the right consumption-output pair.

Definition: Admissible allocations

An allocation is resource-feasible provided that

$$E[c(\omega)] \leq E[y(\omega)] \ ,$$

where the expectation operator indicates that we compute a population average. An allocation is said to be admissible if it is resource feasible and incentive compatible.

Revelation Principle I

Why should we be interested in such artificial communication games?

One can prove a *Revelation Principle* which states the following: An allocation can be reached as the equilibrium outcome of some game if and only if it is incentive compatible, i.e. if and only if it can be reached as the truth-telling equilibrium of a direct mechanism.

This Revelation Principle has to take account of the fact that we are dealing with a large economy, and therefore differ from the typical treatment in textbooks such as e.g. Mas-Colell et al. (1995).

Revelation Principle II

In the following we develop such a revelation principle. This material is admittedly somewhat abstract, and will not be needed in what follows. We nevertheless include into the slides, because it provides the deep theoretical explanation for our interest in incentive-compatible allocations.

Revelation Principle III

An anonymous mechanism $M = [R, c^M, y^M]$ consists of an abstract set of feasible reports R, and two outcome functions $c^M : R \times \Delta(R) \to \mathbb{R}_+$, and $y^M : R \times \Delta(R) \to \mathbb{R}_+$, which specify, respectively, an individuals' consumption level and output requirement as a function of the individuals' own report $r \in R$, and as a function of the cross-section distribution of reports $\rho \in \Delta(R)$.

Consider the game induced by mechanism M in the given economic environment. Suppose that individuals behave according to a strategy $s : \Omega \to M$. We denote by $\rho(s)$ the corresponding cross-section distribution of reports.

Revelation Principle IV

A strategy $s^* : \Omega \to M$ is a *Bayes-Nash-equilibrium (BNE)* if, for all ω, and for all $r \in R$

$$u(c^M(s^*(\omega),\rho(s^*)),y^M(s^*(\omega),\rho(s^*)),\omega) \geq u(c^M(r,\rho(s^*)),y^M(r,\rho(s^*)),\omega) \ .$$

A strategy $s^* : \Omega \to M$ is a *Dominant-Strategy-equilibrium (DSE)* if, for all ω, for all $\rho \in \Delta(R)$ and for all $r \in R$

$$u(c^M(s^*(\omega),\rho),y^M(s^*(\omega),\rho),\omega) \geq u(c^M(r,\rho),y^M(r,\rho),\omega) \ .$$

Revelation Principle V

An allocation $c : \Omega \to \mathbb{R}_+$, $y : \Omega \to \mathbb{R}_+$ is said to be implementable as a *BNE* (*DSE*) if there is a mechanism M with a *BNE* (*DSE*) strategy s^* so that, for all ω,

$$c(\omega) = c^M(s^*(\omega), \rho(s^*)) \quad \text{and} \quad y(\omega) = y^M(s^*(\omega), \rho(s^*)) \; .$$

Theorem 1: Revelation Principle

An allocation is implementable as a *BNE* if and only if it is incentive-compatible.

Taxation Principle I

Definition: Decentralizability of an allocation with an income tax

An allocation $c : \Omega \to \mathbb{R}_+$, $y : \Omega \to \mathbb{R}_+$ can be reached via income taxation if there is a function $\tau : \mathbb{R}_+ \to \mathbb{R}$ such that, for all $\omega \in \Omega$,

$$(c(\omega), y(\omega)) \quad \in \quad \text{argmax}_{c', y' \in \mathbb{R}_+^2} \quad u(c', y', \omega) \quad s.t. \quad c' \leq y' - \tau(y') \; .$$

and

$$E[\tau(y(\omega))] \geq 0 \; .$$

Taxation Principle II

Theorem 2: Taxation Principle

Suppose that private goods consumption is essential, i.e. for all ω, y and y', $u(0, y, \omega) < u(c, y', \omega)$ if $c > 0$. Say that an allocation $c : \Omega \to \mathbb{R}_+$, $y : \Omega \to \mathbb{R}_+$ is reasonable if $c(\omega) > 0$, for all ω.

A reasonable allocation is admissible if and only if it is decentralizable with an income tax.

Taxation Principle III

On the significance of the taxation principle:

We can formulate a problem of optimal income taxation as an optimization over the set of admissible allocations. Once we have found an optimal admissible allocation, we know that there exists an income tax schedule that reaches this outcome.

Useful references for the revelation principle and the taxation principle are Hammond (1979) and Guesnerie (1995).

Taxation Principle IV

The next question then is how to find this tax system. Often there is a degree of freedom in the specification of the tax system. Still, one typically refers $y(\omega) - c(\omega)$ as the implicitly defined tax payment of a type ω-individual, and to

$$1 + \frac{u_2(c(\omega), y(\omega), \omega)}{u_1(c(\omega), y(\omega), \omega)}$$

as the wedge or marginal tax rate for a type ω-individual.

Note that, for a given allocation, these quantities are unambiguously pinned down.

This terminology is explained in what follows.

Taxation Principle V

Tax Payment: Suppose that an allocation $c : \Omega \to \mathbb{R}_+$, $y : \Omega \to \mathbb{R}_+$, is decentralized by means of a tax function τ. Denote by $(c^*(\omega, \tau), y^*(\omega, \tau))$ a type ω-individual's solution to the problem

$$\max_{c', y' \in \mathbb{R}^2_+} \quad u(c', y', \omega) \quad s.t. \quad c' \leq y' - \tau(y') \ .$$

If the tax function τ reaches the allocation $c : \Omega \to \mathbb{R}_+$, $y : \Omega \to \mathbb{R}_+$, then it has to be the case that $c(\omega) = c^*(\omega, \tau)$ and $y(\omega) = y^*(\omega, \tau)$, and hence

$$\tau(y^*(\omega, \tau)) = \tau(y(\omega)) = y(\omega) - c(\omega) \ .$$

Taxation Principle VI

Wedges: The wedge describes the gap between an individuals' marginal rate of substitution between consumption and output provision, which equals

$$-\frac{u_2(c(\omega), y(\omega), \omega)}{u_1(c(\omega), y(\omega), \omega)}$$

and the economy's marginal rate of transformation between output and private goods consumption which equals 1.

Thus the wedge is defined as

$$1 + \frac{u_2(c(\omega), y(\omega), \omega)}{u_1(c(\omega), y(\omega), \omega)} \ .$$

Taxation Principle VII

The wedge is a local measure of the distortion, i.e. of the difference to an allocation that is Pareto-efficient in the set of resource-feasible allocations. Remember, that such first-best allocations always have wedges of 0.

Taxation Principle VIII

Marginal tax rates: The definition of the marginal tax rate assumes that an allocation is decentralized by means of a differentiable tax function τ and that a type ω-individual's solution $(c^*(\omega,\tau), y^*(\omega,\tau))$ to the problem

$$\max_{c',y' \in \mathbb{R}^2_+} \quad u(c',y',\omega) \quad s.t. \quad c' \leq y' - \tau(y') \; .$$

satisfies the first-order condition

$$1 - \tau'(y^*(\omega,\tau)) = -\frac{u_2(c^*(\omega,\tau), y^*(\omega,\tau), \omega)}{u_1(c^*(\omega,\tau), y^*(\omega,\tau), \omega)} \; .$$

Taxation Principle IX

Now, if the tax function τ reaches the allocation $c : \Omega \to \mathbb{R}_+$, $y : \Omega \to \mathbb{R}_+$, then it has to be the case that $c(\omega) = c^*(\omega, \tau)$ and $y(\omega) = y^*(\omega, \tau)$, and hence

$$1 - \tau'(y(\omega)) = -\frac{u_2(c(\omega), y(\omega), \omega)}{u_1(c(\omega), y(\omega), \omega)} .$$

Note that the marginal tax rate is equal to the wedge.

Optimal income taxation and social insurance

We can now state the problem of optimal income taxation in the following way: Choose an allocation $c : \Omega \to \mathbb{R}_+$, $y : \Omega \to \mathbb{R}_+$ so as to maximize a social welfare function

$$S = E[g(\omega)u(c(\omega), y(\omega), \omega)] \ ;$$

over the set of admissible allocations.

Here, $g : \Omega \to \mathbb{R}_+$ is a function that specifies the welfare weights of different types of individuals. We can without loss of generality assume that $E[g(\omega)] = 1$.

Note that we engage here in an interpersonal comparison of utilities, which is a thorny issue.

One way out is to let $g(\omega) = 1$, for all ω and to reinterpret the objective function as the expected utility of one representative individual – so that no interpersonal comparison of utilities is needed – who is at an ex ante stage/ behind a veil of ignorance and will learn her type tomorrow. Thereby we reinterpret the problem of optimal redistributive taxation – where we tax some people with the intention to give the revenue to other people – as a problem of social insurance, i.e. an individual may be willing to give up some consumption in the event of being high-skilled in exchange for increased consumption in the case of being low-skilled.

Two Types: Introduction I

We explore an optimal income tax under the assumption that there are only two types of individuals.

We assume that $\Omega = \{\omega_1, \omega_2\}$, with $\omega_2 > \omega_1$. We denote the population shares of these types by f_1 and f_2, respectively. For simplicity, we assume in this section that preferences are additively separable so that

$$u(c, y, \omega) = v(c) - k\left(\frac{y}{\omega}\right) \ ,$$

where v is an increasing and concave function and and k is an increasing and convex function.

Two Types: Introduction II

The two-type model has a number of interesting features:

- One can solve it using standard tools such as the Kuhn-Tucker approach to constrained optimization problems.
- One can see in a transparent way that wedges/ distortions/ marginal taxes different from zero are part of an optimal tax policy.

The disadvantage however is that the resulting income tax schedule looks somewhat artificial because there are only two income levels.

The following Propositions cover the essential insights of the two-type model.

Two Types: Introduction III

The following References may be useful when you try to prove these Propositions: Stiglitz (1982), Weymark (1987), Hellwig (2007), Bierbrauer and Boyer (2014).

Optimal utilitarian income taxation with two types

Proposition 1

Suppose that v is strictly increasing and strictly concave function with $lim_{c \to 0} v'(c) = \infty$ and $lim_{c \to \infty} v'(c) = 0$, and that k is strictly increasing and strictly convex function with $lim_{h \to 0} k'(h) = 0$ and $lim_{h \to \infty} k'(h) = \infty$. Finally, suppose that the welfare function is such that $g(\omega_2) = g(\omega_1) = 1$.

Proposition 1 ctd

The allocation c^*, y^* which maximizes welfare over the set of resource-feasible allocations has the following properties:

i) $c^*(\omega_1) = c^*(\omega_2)$ and $y^*(\omega_2) > y^*(\omega_1)$.

ii) The incentive constraint of high-skilled individuals is violated, i.e.

$$v(c^*(\omega_2)) - k\left(\frac{y^*(\omega_2)}{\omega_2}\right) < v(c^*(\omega_1)) - k\left(\frac{y^*(\omega_1)}{\omega_2}\right) .$$

iii) The marginal income tax rate for low-skilled individuals is equal to zero.

iv) The marginal income tax rate for high-skilled individuals is equal to zero.

Prop 1 ctd

The allocation c^{**}, y^{**} which maximizes welfare over the set of admissible allocations has the following properties:

i) Monotonicity: $c^{**}(\omega_2) > c^{**}(\omega_1)$ and $y^{**}(\omega_2) > y^{**}(\omega_1)$.

ii) The incentive constraint of low-skilled individuals is not binding, i.e.

$$v(c^{**}(\omega_1)) - k\left(\frac{y^{**}(\omega_1)}{\omega_1}\right) > v(c^{**}(\omega_2)) - k\left(\frac{y^{**}(\omega_2)}{\omega_1}\right) .$$

iii) The incentive constraint of high-skilled individuals is binding, i.e.

$$v(c^{**}(\omega_2)) - k\left(\frac{y^{**}(\omega_2)}{\omega_2}\right) = v(c^{**}(\omega_1)) - k\left(\frac{y^{**}(\omega_1)}{\omega_2}\right) .$$

iv) The marginal income tax rate for low-skilled individuals is positive and is equal to zero for high-skilled individuals.

Proposition 2

Suppose that $v(c) = c$ and that k is strictly increasing and strictly convex function with $lim_{h \to 0} k'(h) = 0$ and $lim_{h \to \infty} k'(h) = \infty$. Finally, suppose that the welfare function is such that $g(\omega_2) = g(\omega_1) = 1$.

There is an allocation c^*, y^* with the following properties:

i) It maximizes welfare over the set of resource-feasible allocations.

ii) The marginal tax rates for high-skilled and for low-skilled individuals are equal to zero.

iii) It it is admissible, i.e. it maximizes welfare also over the set of resource-feasible and incentive-compatible allocations.

Proposition 3

Suppose that $v(c) = c$ and that k is strictly increasing and strictly convex function with $lim_{h \to 0} k'(h) = 0$ and $lim_{h \to \infty} k'(h) = \infty$. Finally, suppose that the welfare function is such that $g(\omega_2) < g(\omega_1)$. The allocation c^{**}, y^{**} which maximizes welfare over the set of admissible allocations has the following properties:

i) Monotonicity: $c^{**}(\omega_2) > c^{**}(\omega_1)$ and $y^{**}(\omega_2) > y^{**}(\omega_1)$.

ii) The incentive constraint of low-skilled individuals is not binding.

iii) The incentive constraint of high-skilled individuals is binding.

iv) The marginal income tax rate for low-skilled individuals is positive.

v) The marginal income tax rate for high-skilled individuals is equal to zero.

Bierbrauer and Boyer (2014): More details I

- Complete characterization of Pareto frontier.
- Continuum of individuals of mass 1 with quasi-linear preferences in leisure:

$$U_i = u(c_i) - l_i$$

- Standard two-type model of optimal income taxation: $0 < w_L < w_H$ and $f_H = 1 - f_L$.
- Simplification: non-negativity constraints on individual consumption levels and on income assumed away

Bierbrauer and Boyer (2014): More details II

Optimization problem:

$$V_H(v_L, r) \quad := \quad max \quad u(c_H) - \tfrac{y_H}{w_H} \quad \text{subject to}$$

$$f_H\, y_H + f_L\, y_L = r + f_H\, c_H + f_L\, c_L\ ,$$

$$u(c_H) - \frac{y_H}{w_H} \geq u(c_L) - \frac{y_L}{w_H}\ ,$$

and

$$u(c_L) - \frac{y_L}{w_L} \geq u(c_H) - \frac{y_H}{w_L}\ ,$$

$$u(c_L) - \frac{y_L}{w_L} = v_L\ .$$

Proposition 1 in the paper

For given r, the function V_H has the following properties:

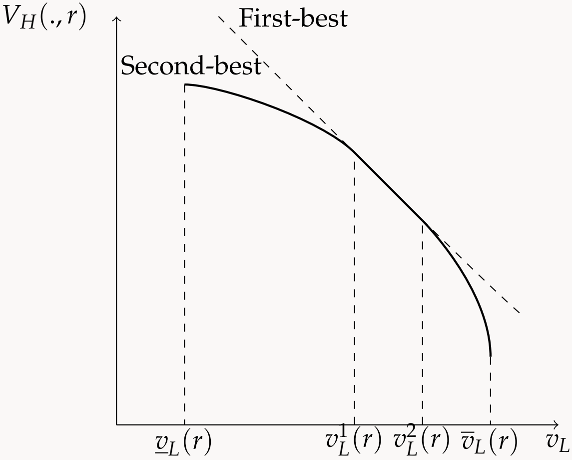


Figure: Pareto-Frontiers.

Proposition 2 in the paper

Both marginal tax rates are non-decreasing functions of v_L. We also have that $\tau_H(v_L, r) \leq 0$ and that $\tau_L(v_L, r) \geq 0$, for all v_L and r. More specifically,

(a) For $v_L \in [\underline{v}_L(r), v_L^1(r)[$, $\tau_H < 0$, $\tau_{H1} > 0$, and $\tau_L = 0$.

(b) For $v_L \in [v_L^1(r), v_L^2(r)]$, $\tau_H = 0$, and $\tau_L = 0$.

(c) For $v_L \in]v_L^2(r), \overline{v}_L(r)]$, $\tau_H = 0$, $\tau_L > 0$, and $\tau_{L1} > 0$.

Optimal income taxation with preferences that are quasilinear in consumption: Motivation

Since Diamond (1998) the setup with preferences that are quasilinear in consumption has received considerable attention.

One reason is that this gives us an easy to interpret formula of the determinants of optimal marginal tax rates, which is known as Diamond's *ABC*-formula.

Optimal income taxation with preferences that are quasilinear in consumption: Motivation

Another reason is that this setup makes it particularly easy to calibrate an optimal tax schedule. We will illustrate this later.

Another feature of this setup is that the optimal income tax/ the optimal incentive compatible allocation can be characterized with no need to introduce new mathematical tools such as the theory of optimal control or the calculus of variations.

Setup

Productivity levels are drawn from a compact interval $\Omega = [\underline{\omega}, \overline{\omega}] \subset \mathbb{R}_+$.

Preferences are quasilinear in consumption so that

$$u(c(\omega), y(\omega), \omega) = c(\omega) - k(y(\omega), \omega) \ ,$$

where k is a cost function with the following properties:

- For any ω, $k_1(\cdot) > 0$ and $k_{11}(\cdot) > 0$.
- For any ω, $lim_{y' \to 0} \, k_1(y, \omega) = 0$, and $lim_{y' \to \infty} \, k_1(y, \omega) = \infty$.
- For any $y > 0$ and any $\omega \in (\underline{\omega}, \overline{\omega})$, $k_{12}(\cdot) < 0$.

The objective is to maximize a welfare function

$$S = E[g(\omega)(c(\omega) - k(y(\omega), \omega))] \ ,$$

where g is a strictly decreasing function.

Characterization of incentive-compatible allocations

Lemma 1

Assume that $c : \Omega \to \mathbb{R}_+$ and $y : \Omega \to \mathbb{R}_+$ are twice differentiable.

Denote: $U(\omega) := c(\omega) - k(y(\omega), \omega)$.

Then: (c, y) is incentive compatible if and only if the following two conditions are met:

i) *Envelope Condition:* For all $\omega \in (\underline{\omega}, \overline{\omega})$, $U'(\omega) = -k_2(y(\omega), \omega)$.

ii) *Monotonicity:* For all $\omega \in (\underline{\omega}, \overline{\omega})$, $y'(\omega) \geq 0$.

Characterization of admissible allocations I

Lemma 2

Consider a pair consisting of a number $U(\underline{\omega})$ and a function $y : \Omega \to \mathbb{R}_+$ that is twice differentiable.

There is $c : \Omega \to \mathbb{R}_+$ so that (c, y) is admissible if and only if the following two conditions are met:

i) *Envelope Condition meets Resource Constraint:*

$$U(\underline{\omega}) \leq E\left[y(\omega) - k(y(\omega), \omega) + \frac{1 - F(\omega)}{f(\omega)} k_2(y(\omega), \omega)\right] . \quad (1)$$

ii) *Monotonicity:* For all $\omega \in (\underline{\omega}, \overline{\omega})$, $y'(\omega) \geq 0$.

Characterization of admissible allocations II

Significance of the Lemma:

Can characterize an admissible allocation via a number $U(\underline{w})$ and a monotonic function $y : \Omega \to \mathbb{R}_+$.

The indirect utility function $U : \Omega \to \mathbb{R}$ is then characterized via

$$U(\omega) = U(\underline{\omega}) + \int_{\underline{\omega}}^{\omega} U'(s)\, ds$$

$$= U(\underline{\omega}) - \int_{\underline{\omega}}^{\omega} k_2(y(s), s)\, ds\ .$$

The function $c : \Omega \to \mathbb{R}_+$ is then characterized via

$$c(\omega) = U(\omega) + k(y(\omega), \omega)$$

$$= U(\underline{\omega}) - \int_{\underline{\omega}}^{\omega} k_2(y(s), s)\, ds + k(y(\omega), \omega)\ .$$

Characterization of admissible allocations III

The resource constraint can then be rewritten as follows:

After an integration by parts, we find that

$$E[c(\omega)] = U(\underline{\omega}) + E\left[k(y(\omega),\omega) - \tfrac{1-F(\omega)}{f(\omega)} k_2(y(\omega),\omega)\right] .$$

Substituting this into $E[c(\omega)] \leq E[y(\omega)]$ yields (1).

Welfare induced by an admissible allocation I

Let (c, y) be characterized by a pair $(U(\underline{w}), y)$.

Then, welfare is given by

$$\begin{aligned} S &= E[g(\omega)U(\omega)] \\ &= U(\underline{\omega}) + E\left[g(\omega)\left(\int_{\underline{\omega}}^{\omega} U'(s)ds\right)\right] \\ &= U(\underline{\omega}) - E\left[g(\omega)\left(\int_{\underline{\omega}}^{\omega} k_2(y(s), s)ds\right)\right] . \end{aligned}$$

Welfare induced by an admissible allocation II

After an integration by parts, and upon using that $E[g(\omega)] = 1$, this can be rewritten as

$$S = U(\underline{\omega}) - E\left[G(\omega)\frac{1-F(\omega)}{f(\omega)}k_2(y(\omega),\omega)\right] , \tag{2}$$

where

$$G(\omega) := E[g(s) \mid s \geq \omega] = \int_{\omega}^{\overline{\omega}} g(s)\frac{f(s)}{1-F(\omega)}\, ds .$$

Welfare induced by an admissible allocation III

For any allocation that is Pareto-efficient in the set of admissible allocations, the resource constraint will bind so that, by (1),

$$U(\underline{\omega}) = E\left[y(\omega) - k(y(\omega),\omega) + \frac{1-F(\omega)}{f(\omega)}k_2(y(\omega),\omega)\right] .$$

Hence, once we have found the optimal y-function, this equation gives us the optimal value of $U(\underline{\omega})$.

Upon substituting this expression into the objective function, we obtain

$$S = E\left[y(\omega) - k(y(\omega),\omega) + (1-G(\omega))\tfrac{1-F(\omega)}{f(\omega)}k_2(y(\omega),\omega)\right] \quad (3)$$

Note that this function solely depends on y.

Welfare induced by an admissible allocation IV

There are now two optimization problems of interest:

- **The full problem:** Choose y so as to maximize the expression in (3) subject to the constraint that $y'(\omega) \geq 0$, for all $\omega \in (\underline{\omega}, \overline{\omega})$.
- **The relaxed problem:** Choose y so as to maximize the expression in (3).

Obviously, if the solution to the relaxed problem is monotonic, then it is also a solution to the full problem.

Solving the relaxed problem

Proposition 5

A solution to the relaxed problem has the following properties:

1. For any ω, $y(\omega)$ satisfies the following equation:

$$\tau'(y(\omega)) = 1 - k_1(y(\omega), \omega) = -(1 - G(\omega))\frac{1 - F(\omega)}{f(\omega)} k_{12}(y(\omega), \omega) \; .$$

2. The marginal tax rate is strictly positive for all $\omega < \overline{\omega}$.
3. The marginal tax rate is zero for $\omega = \overline{\omega}$.

The relaxed problem and the full problem I

Proposition 6

Suppose that $k(y(\omega), \omega) = \left(\frac{y(\omega)}{\omega}\right)^{1+\frac{1}{\varepsilon}}$, for some $\varepsilon > 0$. Then, the solution to the relaxed problem can be written as

$$\frac{\tau'(y(\omega))}{1-\tau'(y(\omega))} = \frac{1-k_1(y(\omega),\omega)}{k_1(y(\omega),\omega)} = (1-G(\omega))\,\frac{1-F(\omega)}{f(\omega)\,\omega}\left(1+\frac{1}{\varepsilon}\right) .$$

Consider an interval $\Omega' = [\omega_a, \omega_b]$ over which g is constant. Suppose that h with $h(\omega) := \frac{f(\omega)\omega}{1-F(\omega)}$ is an increasing function, then the solution to the relaxed problem satisfies $y'(\omega) \geq 0$, for all $\omega \in \Omega'$.

The relaxed problem and the full problem II

Interpretation of Proposition 6

- The equation is known as Diamond's *ABC*-formula.
- The left-hand-side is an increasing function of the marginal tax rate.
- Thus the marginal tax rate at a certain level of income is higher if

  A: we care less for individuals who earn this level of income or more,

  B: the hazard rate is lower,

  C: the elasticity ε is lower.
- Note that ε is the elasticity of labor supply with respect to the net wage rate.

Quantitative implications of the *ABC*-formula: Pareto-distribution

Pareto-Distribution:

- Continuous distribution with support $[x_0, \infty[$.
- *CDF* given by $F = 1 - \left(\frac{x_0}{x}\right)^a$.
- Hazard rate constant, thus $\frac{1-F(x)}{f(x)} \frac{1}{x} = \frac{1}{a}$.

Formula in Proposition 6 simplifies to

$$\frac{\tau'(y(\omega))}{1-\tau'(y(\omega))} = \frac{1-k_1(y(\omega),\omega)}{k_1(y(\omega),\omega)} = (1-G(\omega))\,\frac{1}{a}\left(1+\frac{1}{\varepsilon}\right) .$$

Note that, with a Pareto-distribution, there is no top.

Wage and Income Date for the US, 1992 taken from Diamond (1998)

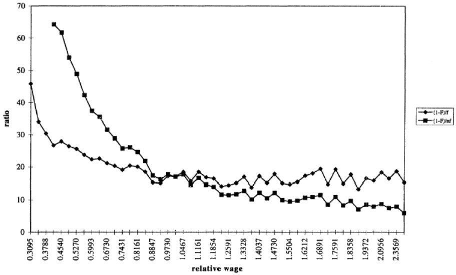
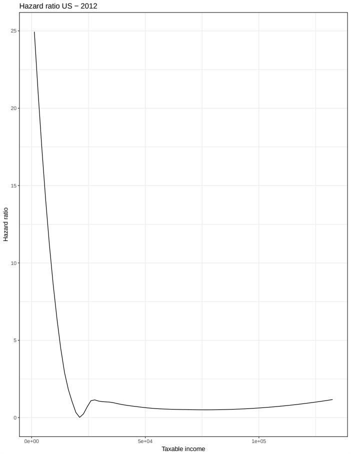
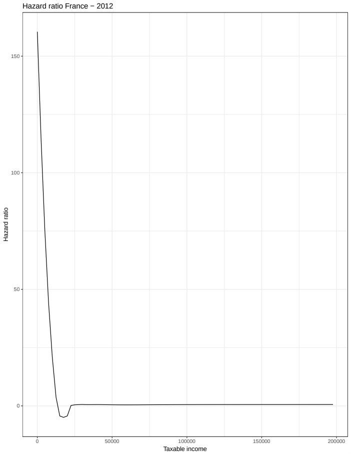


FIGURE 1. RATIOS $[1 - F(n)]/f(n)$ AND $[1 - F(n)]/[nf(n)]$ CALCULATED FROM RELATIVE WAGES

Suggests: Pareto-Distribution is plausible for high incomes.





See also Blanchet, Fournier and Piketty (2017).

Top marginal tax rates, taken from Diamond (1998)

Suppose that $G(\omega)$ converges to a number g as ω goes to ∞. Then the "top" marginal tax rate is given by

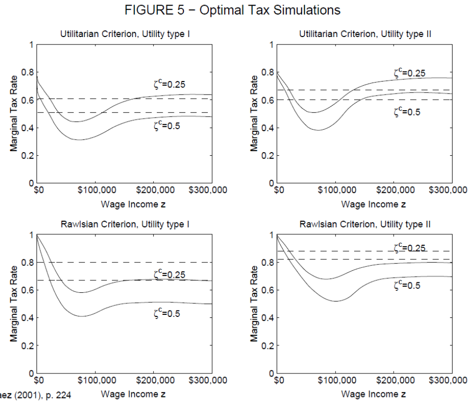
$$\frac{\tau(y_\infty)}{1-\tau(y_\infty)} \simeq (1-g)\ \frac{1}{a}\left(1+\frac{1}{\varepsilon}\right)\ .$$

TABLE 1—ASYMPTOTIC MARGINAL TAX RATES

	$g = 0$			$g = 0.25$			$g = 0.5$		
$a =$	0.5	1.5	5.0	0.5	1.5	5.0	0.5	1.5	5.0
e									
0.2	92	80	55	90	75	47	86	67	38
0.5	86	67	38	82	60	31	75	50	23

Notes: Asymptotic marginal tax rates, in percent, with a constant elasticity of labor supply, e, a Pareto distribution of skills with parameter a, and a ratio of social marginal utility with infinite income to average social marginal utility of g.

U-shape schedule of marginal tax rates



Source: Saez (2001), p. 224

Controversy between

Mankiw, N. Gregory, Matthew C. Weinzierl, and Danny Yagan. 2009. Journal of Economic Perspectives.

Diamond, Peter, and Emmanuel Saez. 2011. Journal of Economic Perspectives.

Mankiw et al. argue that one could also approximate the income distribution with a lognormal distribution, and that this implies that the optimal tax schedule is approximately flat for high levels of income.

Diamond and Saez argue that the Pareto-distribution is the "right" distribution.

Public Finance

Pierre Boyer

École polytechnique - CREST

Fall 2025