Tax Policy and Inequality

Optimal Taxation

Damon Jones

Harris School of Public Policy
University of Chicago

Outline

Taxes and Economic Activity

- Reduce rewards to work, savings, investment?
- Reallocate activity across different sectors/goods?
- Divert resources to compliance and evasion?
- Redistribute well-being across individuals?

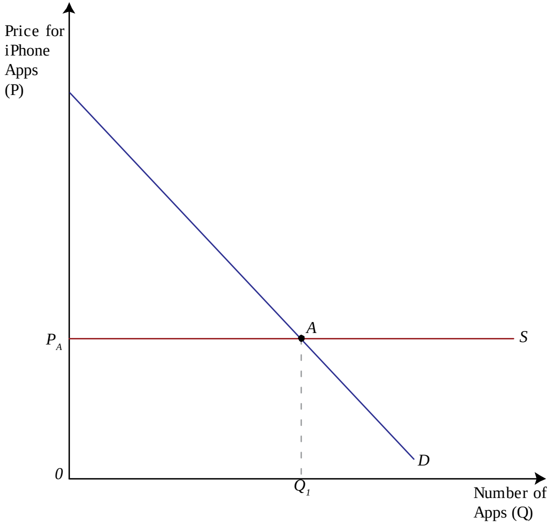
Cost: Marginal Cost of Public Funds

- Suppose we collect \$1 in tax revenue
- Cost of raising this revenue is more than \$1
  - Referred to as deadweight loss
  - Excess burden
- Estimates vary: e.g. 30¢

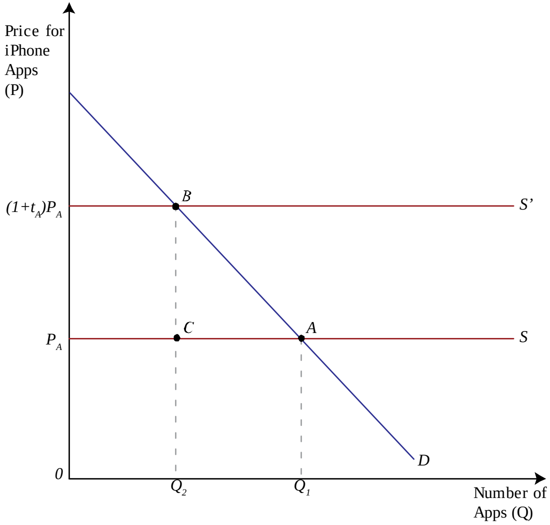
Excess Burden: Graphical Analysis

- We can measure excess burden with (compensated) demand and supply curves
  - this is the standard dead weight loss that we are used to
- We will consider a simple example:
  - Assume constant marginal costs of providing iPhone Apps
  - Consider an ad valorem tax of t_A levied on Apps

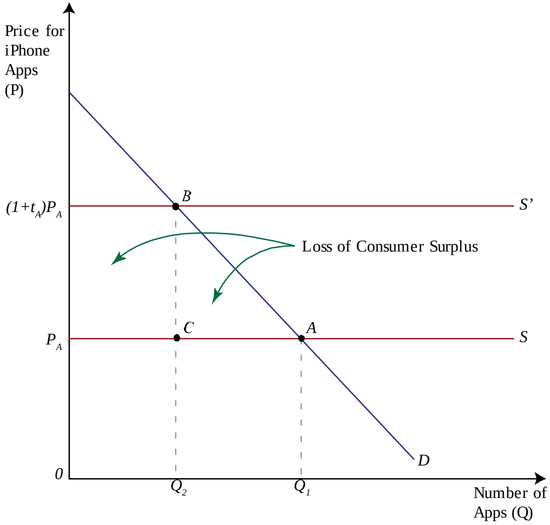
Excess Burden: Graphical Analysis



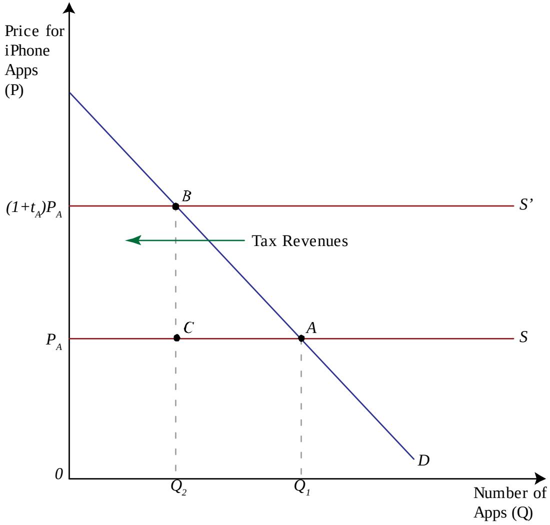
Excess Burden: Graphical Analysis



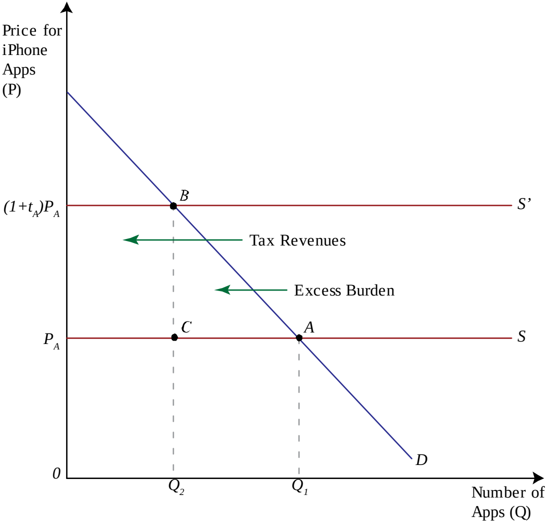
Excess Burden: Graphical Analysis



Excess Burden: Graphical Analysis



Excess Burden: Graphical Analysis



Excess Burden: Graphical Analysis

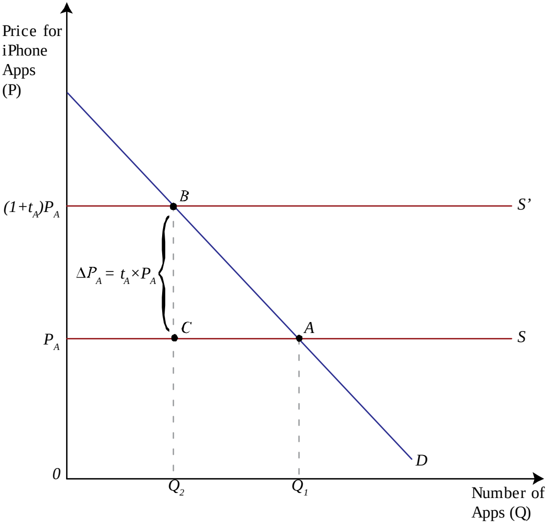
- The excess burden is the triangle ABC
  - What is the area of this triangle?

$$
\begin{aligned}
Area &= \frac{1}{2} \times \text{base} \times \text{height} \\
&= \frac{1}{2} \times \triangle Q \times \triangle P_A
\end{aligned}
$$

- First take $\triangle P_A$:

$$
\begin{aligned}
\triangle P_A &= (1 + t_A)\, P_A - P_A \\
&= P_A + t_A \times P_A - P_A \\
&= t_A \times P_A
\end{aligned}
$$

Excess Burden: Graphical Analysis



Excess Burden: Graphical Analysis

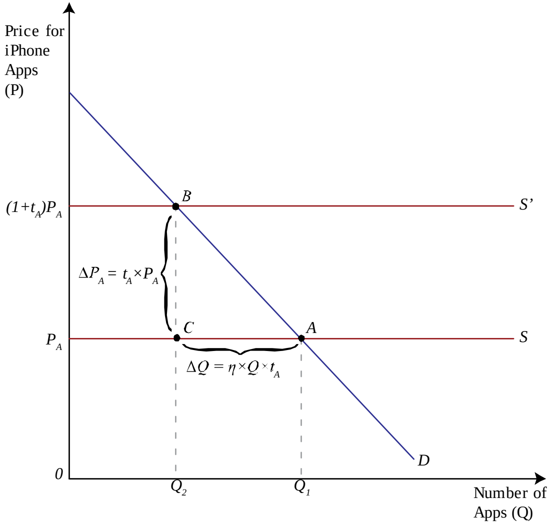
- Now consider $\triangle Q$
  - Use the definition of $\eta =$ elasticity of (compensated) demand:

$$\begin{aligned} \eta &= \frac{\triangle Q}{\triangle P_A} \frac{P_A}{Q} \\ \triangle Q &= \eta \left(\frac{Q}{P_A} \right) \triangle P_A \end{aligned}$$

  - We already showed that $\triangle P_A = t_a \times P_A$, so:

$$\begin{aligned} \triangle Q &= \eta \left(\frac{Q}{P_A} \right) t_A P_A \\ &= \eta \times Q \times t_A \end{aligned}$$

Excess Burden: Graphical Analysis



Excess Burden: Graphical Analysis

- Putting the two together, we get:

$$\begin{aligned} \text{Excess Burden} &= \frac{1}{2} \times (\triangle P_A) \times (\triangle Q) \\ &= \frac{1}{2} \times (t_A \times P_A) \times (\eta \times Q \times t_A) \\ &= \frac{1}{2}\eta \times P_A Q \times \left(t_A^2\right) \end{aligned}$$

- Thus, the amount of excess burden depends on:
  1. the sensitivity of demand to price: η
  2. the initial expenditures on the good: $P_A Q$
  3. the square of the tax: t_A^2

Excess Burden: Applied Estimates

- What would be the excess burden of a 10% tax on iPhone Apps?
  - Total App sales in first year: \$213 million
  - Tax rate: 10%
  - Elasticity of demand for apps?
  - Plug in 1.0?
- Excess burden would be approximately:

$$
\begin{aligned}
EB &= \frac{1}{2} \times \eta \times P_A Q \times \left(t_A^2\right) \\
&= \frac{1}{2} \times (1) \times (\$213 \text{ mil}) \times (0.10)^2 \\
&= \$1.065 \text{ million}
\end{aligned}
$$

Beneifit: Revenue and Laffer Curve

- What is the relationship between the tax level and revenue?
- Arthur Laffer $\rightarrow$ Laffer Curve
- Which two tax rates generate zero revenue:?
  - In general there is a revenue maximizing rate
  - Diamond and Saez (2012) derive the maximal rate
  - Estimated bteween 48%-76%

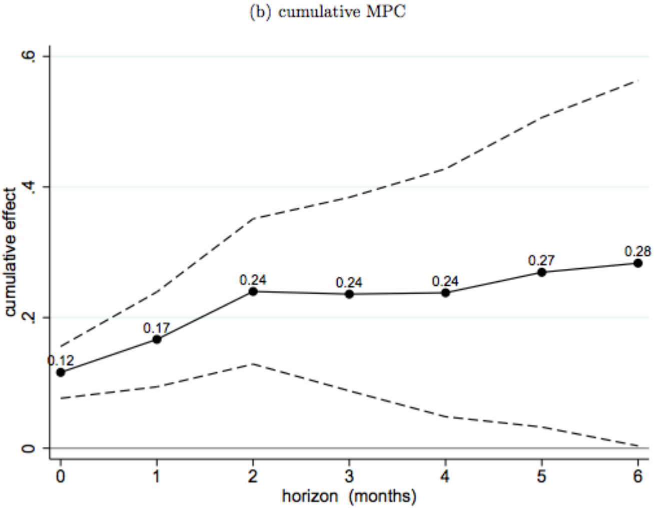
Benefit: Taxes as a Stimulus?

- Keynesian policy/ fiscal policy
  - Tax cuts and spending boost economy/ mitigate recessions
- Discredited in the late 1970s with stagflation
- Revisited since 2001, 2008-2009
  - Government spending $>$ tax cuts
  - Requires valuable government projects

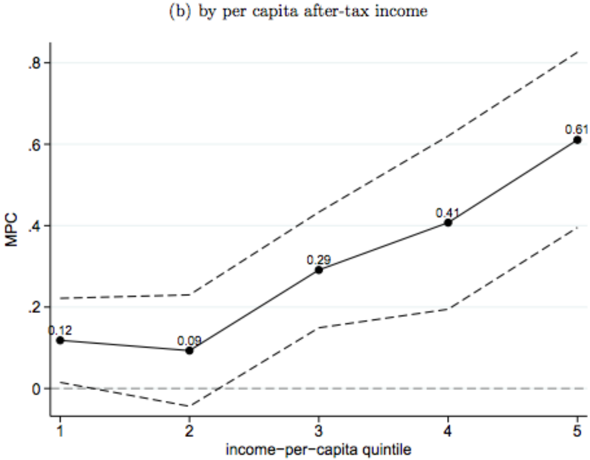
Benefit: Taxes as a Stimulus?

- Depends on whether tax cut is viewed as temporary, permanent, or "very permanent"
- Also depends on the marginal propensity to consume: MPC
- Recent evidence: Lorenz Kueng (2016)
  - Alaska Permanent Fund
  - Average MPC = 30%
  - Largest MPC for higher incomes
  - High MPC for low income, low liquid wealth households

Benefit: Taxes as a Stimulus?



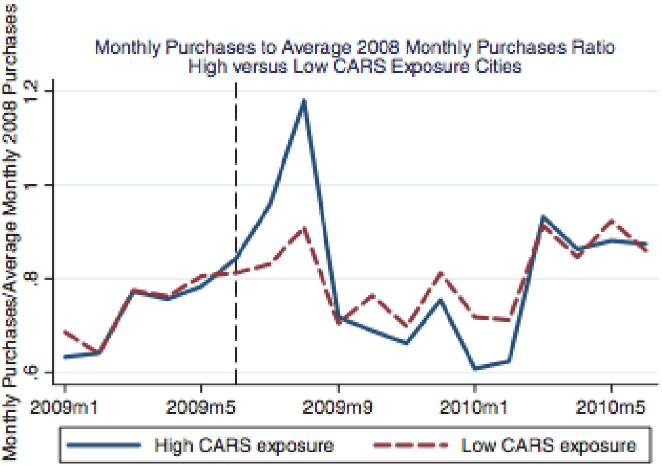
Benefit: Taxes as a Stimulus?



Benefit: Taxes as a Stimulus?

- Alternative: accelerating spending:
  - Cash for clunkers (Mian & Sufi, 2012)
  - Home mortgage interest discounts
- Dismount is important as well
  - Short-run bump up in spending
  - Dip down in the longer run

Benefit: Taxes as a Stimulus?



Benefit: Taxes as a Stimulus?

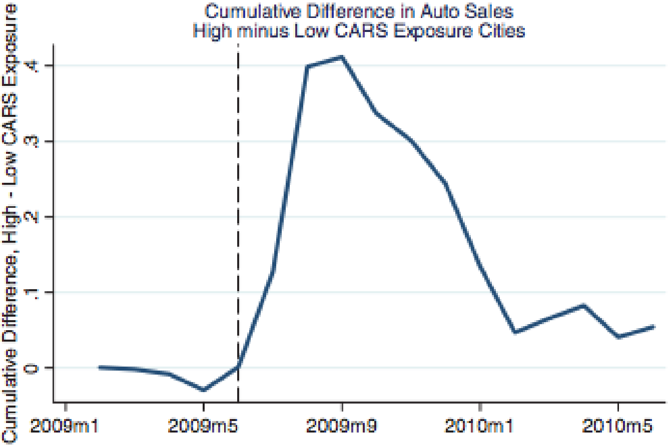


FIGURE IV

Auto Purchases for High and Low CARS Exposure Cities

Benefit: Automatic Stabilization

- Tax schedule is progressive
  - Automatic adjustment in average tax rate as income lowers
- Increase in refundable credits as income drops (EITC)
- Other stabilizers (safety net)
  - Unemployment Income
  - SNAP, etc.

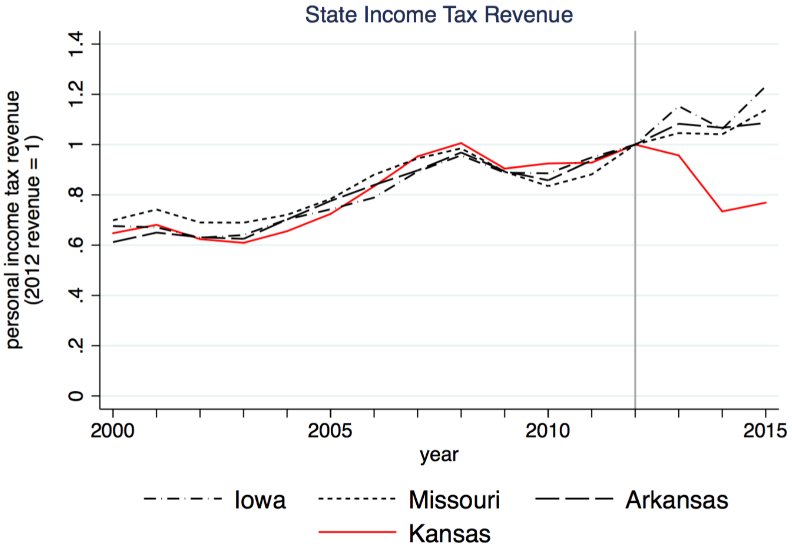
Cost v. Benefit: Optimal Taxation Debate

- Need to compare benefit of taxation to cost
    - Cost includes deadweight loss
- In addition, evaluate redistribution
    - Positive analysis: how much will individuals respond/ who will bear burden?
    - Normative analysis: how do we tradeoff utility across people?
    - Econ: comparative advantage in Pos., not Norm.
- Caution: "Expert" opinions conflate scientific and personal

Cost: Taxes and Growth

- Hard to measure relationship between taxes and growth
  - Only cross country or time series data
- What can we say?
  - Cutting taxes not sufficient: 90%+ MTR 1950s-60s
  - Cutting taxes $\rightarrow$ less revenue (Laffer Curve)

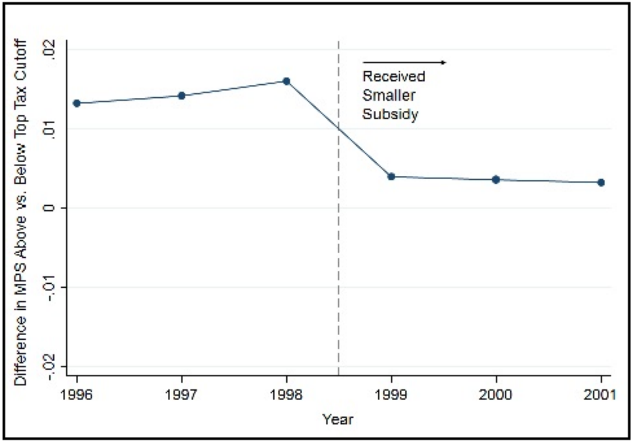
Cost: Taxes and Growth: DeBacker, Heim, Ramnath, and Ross (2017)



Cost: Labor Supply & Savings

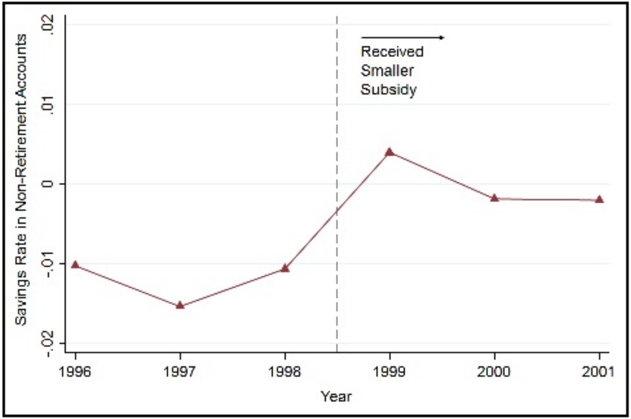
- Historically: little or small effect on labor supply of prime aged, primary earners
  - Larger effect on secondary earners (historically women)
  - Large MTR for secondary earner with high income spouse
- Savings:
  - Mixed evidence on response to subsidies on savings
  - Best Evidence: Chetty, Friedman, Leth-Petersen, Nielsen, Olsen (2014)

Cost: Labor Supply & Savings



When individuals in the top tax bracket received a sm subsidy for retirement savi started saving less in re accounts...

Cost: Labor Supply & Savings



... but the same individuals the amount they were savin retirement accounts by almo the same amount, leavi savings essentially unchan estimate that each $1 of go expenditure on the subsid total savings by 1 cent.

Outline

The Ramsey Rule

- What is the trade-off involved in taxing a good?
  1. The cost of a tax on a good will be the deadweight loss created
  2. The benefit of a tax on a good will be the tax revenue
- The goal is to minimize (1) across taxed goods while allowing the sum of (2) to reach some required amount

The Ramsey Rule

- Let's say the government has N different goods that it may tax
- Formally, the government's problem is:

$$\min_{\{t_i\}} \left(DWL_1 + DWL_2 + \cdots + DWL_N\right)$$
$$s.t.\ R_1 + R_2 + \cdots + R_N = \bar{R}$$

- To solve this problem, we have to set up a Lagrangian:

$$\left(DWL_1 + DWL_2 + \cdots + DWL_N\right) + \lambda\left(\bar{R} - R_1 - R_2 - \cdots - R_N\right)$$

The Ramsey Rule

- The first order condition t_i is :

$$\frac{MDWL_i}{MR_i} = \lambda$$

  - This is referred to as the **Ramsey Rule**
- The ratio of marginal deadweight loss to marginal revenue is the same across goods

$$\frac{MDWL_i}{MR_i} = \frac{MDWL_j}{MR_j}$$

The Ramsey Rule

- Intuitively, consider the following case:

$$\frac{MDWL_i}{MR_i} > \frac{MDWL_j}{MR_j}$$

- If this is the case, we can raise the tax on good j and lower the tax on good i

The Ramsey Rule

- We can interpret the result in terms of elasticities
- First, solve for *MDWL*:

$$DWL = \frac{1}{2}\eta \times PQ \times t^2$$

$$MDWL = \eta \times PQ \times t$$

The Ramsey Rule

- Now, solve for MR:

$$\begin{aligned} R &= t \times PQ \\ MR &= PQ \end{aligned}$$

- Finally, we have:

$$\frac{MDWL}{MR} = \eta t$$

The Ramsey Rule

- Thus, we can rewrite the Ramsey Rule as:

$$\frac{MDWL_i}{MR_i} = \eta_i t_i = \lambda$$

- We can rearrange things:

$$t_i = \frac{\lambda}{\eta_i}$$

- Also, since the marginal dead weight loss rises with the tax rate ($MDWL_i = \eta \times PQ \times t$), we should spread out the tax across a broad base
  - Better to have a 1% tax rate on many goods than a 2% tax rate on a few goods

The Ramsey Rule

- Another way to think about it is to recall the following:

$$\begin{aligned} \triangle P &= tP \text{ and} \\ \eta &= \frac{\triangle Q}{Q}\frac{P}{\triangle P} \\ &= \frac{\triangle Q}{Q}\frac{P}{tP} \\ &= \frac{\triangle Q}{Q}\frac{1}{t} \end{aligned}$$

- Going back to the Ramsey Rule:

$$\begin{aligned} \lambda &= \eta_i t_i \\ &= \left(\frac{\triangle Q_i}{Q_i}\frac{1}{t_i}\right) t_i \\ &= \frac{\triangle Q_i}{Q_i} \end{aligned}$$

The Ramsey Rule

- Thus, yet another way to think about the Ramsey Rule is that the optimal combination of taxes causes an equal *proportional* decrease in quantities:

$$\frac{\triangle Q_i}{Q_i} = \frac{\triangle Q_j}{Q_j}$$

- An important note is that we have thus far ignored the effect of prices changes across markets (i.e. elasticities of substitution)
  - The math becomes messier, but the main results still hold

Equity versus Efficiency

- The standard Ramsey Rule only deals with efficiency
  - What if we had two goods to tax: caviar and cereal
  - Suppose the demand for cereal was much more inelastic
- If caviar is disproportionately consumed by high income individuals, we may place a higher tax than implied by the Ramsey Rule, to increase equity
- Taking equity into account involves two questions:
  - What is the degree to which society desires equity?
  - How different are the tastes of the rich and the poor?

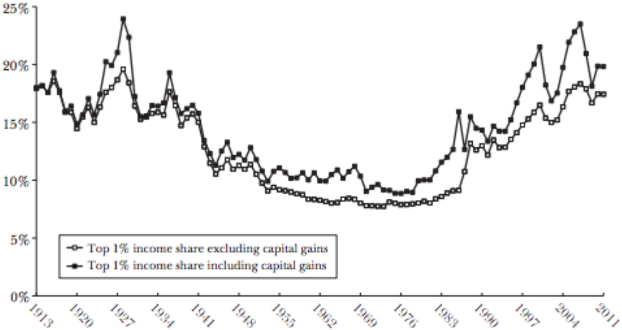
Outline

Share of Income: Top 1% of earners

- Start with Alvaredo, Atkinson, Piketty, and Saez (2013)
- Summary of previous work, including Piketty and Saez (2003) and Piketty, Stancheva, and Saez (2014)
- Use administrative tax data to track top income shares over the 20th century
- Historical analysis and cross-country analysis
- Also consider income + wealth distributions

Share of Income: Top 1% of earners

Figure 1
Top 1 Percent Income Share in the United States



Source: Source is Piketty and Saez (2003) and the World Top Incomes Database.

Share of Income: Top 1% of earners

- Not just technology: patterns differ across similar countries
- Real economic effect of just tax avoidance?
- Behavioral change in effort: should show up in economic growth
- Bargaining between top earners and firms over surplus
- Top income shares negatively correlated with top marginal tax rates

Share of Income: Top 1% of earners

- Wealth/inheritance inequality grew as well, primarily in European countries
- Related to return to capital, relative to economic growth (Piketty)
- Top income and top wealth rankings are correlated (not perfectly)
- The correlation in income and wealth rankings has gotten stronger over time

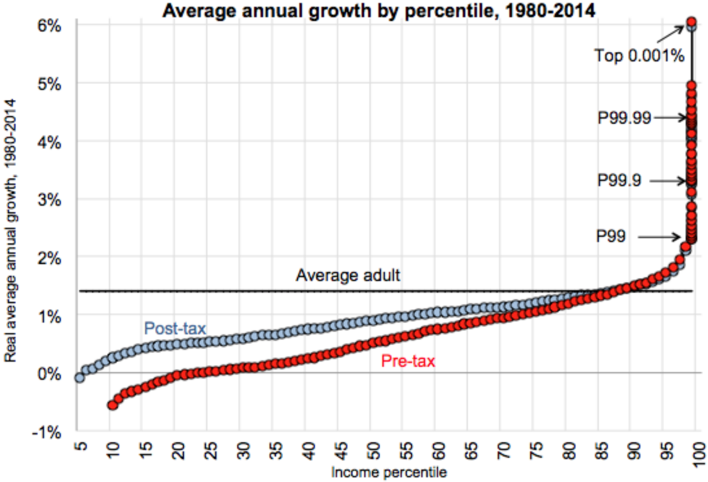
Distributional National Accounts

- Piketty, Saez, and Zucman (2017)
- Gap between micro data-based studies and macro measures of income
- Previous analysis ignored the role of taxes, transfers, and public spending
- Previous studies use the tax unit as the unit of observation: e.g. no ability to separately analysis women and men
- Distributional National Accounts
  - Combine survey, tax, and national accounts data
  - Assigns 100% of national income to individuals
  - Analyze patterns at different pertentiles of the income distribution

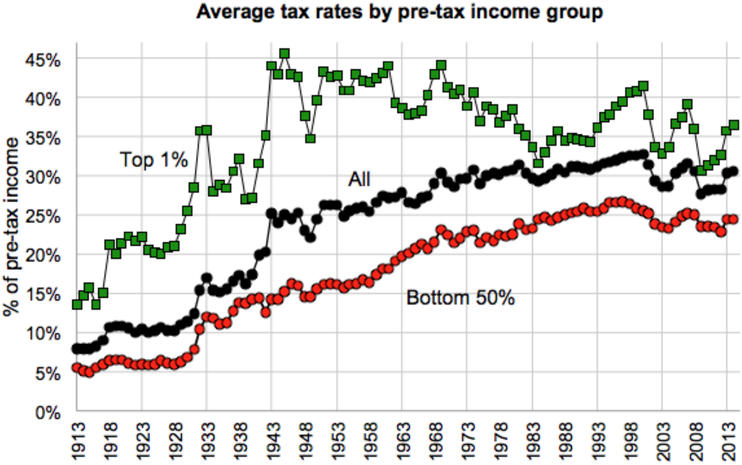
Distributional National Accounts: Methodology

- National income: GDP minus capital depreciation, plus net foreign income
- Three types of income:
  - Factor income: assign national income, labor and capital (includes fringe benefits)
  - Pre-tax income: labor/capital income (tax returns) + pensions, adding back payroll taxes, assign wealth/capital income/corporate profits to individuals, add Social Security, UI, DI
  - Post-tax income: subtract taxes, add individual transfers, distribute government spending
- Requires assumptions about incidence, corporate profits, public goods, government deficits

Distributional National Accounts

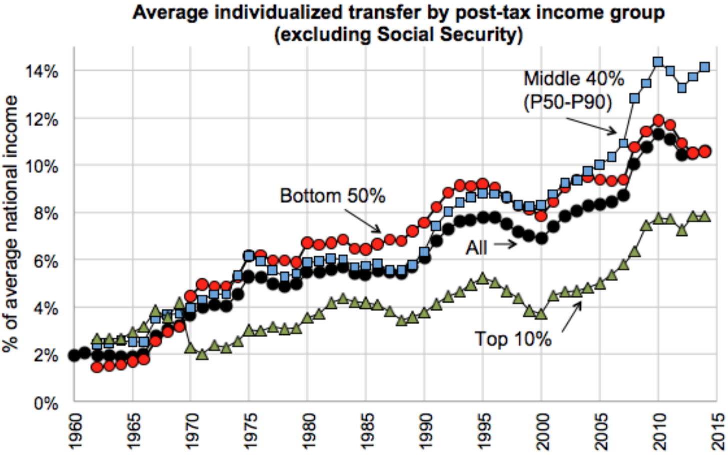


Distributional National Accounts



Source: Appendix Table II-G1.

Distributional National Accounts



Source: Appendix Table II-G4.

Distributional National Accounts: Results

- Pre-tax income share of 1%: 20.2% (15.7% after tax)
- Top 0.1% share close to bottom 50% share
- Middle 40% roughly earns 40% of income
- Tax and transfers generally progressive
- Growth:
  - 1946-1980: Growth more equitable, bottom grew more than top
  - 1980-2014: Bottom 50% stagnant, lower 20% declines in earnings, skewed growth
  - Taxes and transfers moderate growth differences somewhat
  - Closing of gender gaps reduces inequality, but less so for highest incomes
  - Top 1% growth due to wages 1980-1990s, due to capital income late 1990s onward
  - Taxes and transfers have become less progressive (mainly to middle class)

Inequality Overstated?

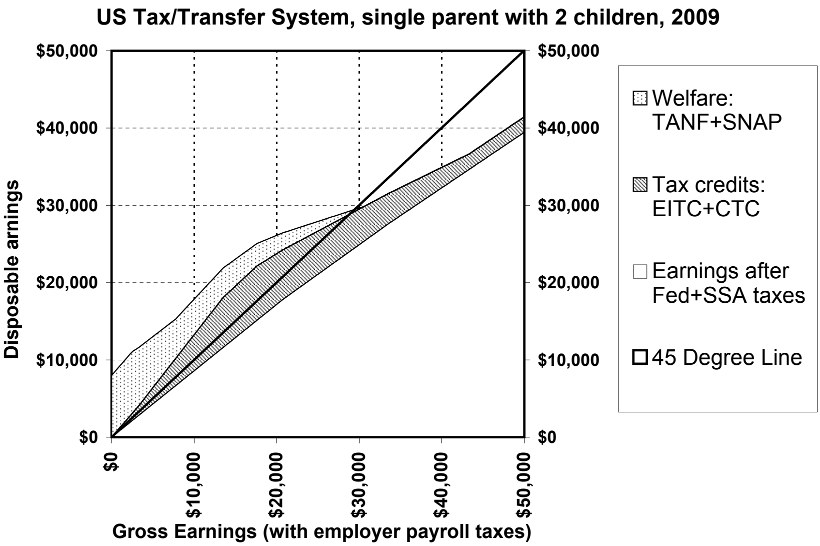
- Auten and Splinter (2017)
- Challenge notion that top 1% income share has doubled over time
- Primary reference: Piketty and Saez (2003)
- Account for non-covered income, tax policy (TRA 1986), demographic change
- Change in top income share goes from 11.2 ppt to 1.7 ppt!
- Rich were rich in 1960s, just hid their money in corporations
- Differences from Piketty, Saez, and Zucman (2017):
  - Treatment of retirement income
  - Underreported income
  - Deficits, dependents, married couples
- Debate as of yet unresolved

T(z)

1. Transfer benefit with zero earnings $-T(0)$ [sometimes called demogrant or lump sum 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:

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

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



Source: Federal Govt

Optimal Income Tax without Behavioral Responses

- Utility $u(c)$ strictly increasing and concave
- $u(c)$ same for everybody where c is after tax income.
- Income is z and **is fixed** for each individual, $c = z - T(z)$
- z has distribution with density $h(z)$
- Government maximizes Utilitarian objective:

$$\max_{T(\cdot)} \int_0^\infty u(z - T(z))h(z)dz$$
$$\text{s.t.} \int_0^\infty T(z)h(z)dz \geq R$$

- Solution: $T(z) \to c = \bar{z} - R$
- 100% marginal tax rate; perfect equalization of after-tax income. Utilitarianism with diminishing marginal utility leads to egalitarianism. With heterogeneity: $u_i'(c) = \mu$

Optimal Income Tax without Behavioral Responses

- **No behavioral responses**: Obvious missing piece: 100% redistribution would destroy incentives to work and thus the assumption that z is exogenous is unrealistic
  - Optimal income tax theory incorporates behavioral responses (Mirrlees REStud '71)
- **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 govt can do [political economy]
- **Heterogeneous Preferences:** Holding $u_i'(c)$ constant means redistributing more towards those with a higher preference for consumption: required health expenses, number of dependent children, or high ability to enjoy consumption

Sufficient Statistic Approach Overview

- Work of Diamond (1998), Piketty (1997) and Saez (2001) bring the Mirrlees (1971) tax formula in line with empirical data
- Build up to general, optimal non-linear tax:
  - Revenue maximizing linear tax
  - Revenue maximizing non-linear tax [Rawlsian SWF]
  - Optimal linear tax
  - Optimal top marginal tax rate
  - Optimal nonlinear tax schedule
- Will sometimes consider case with no income effects for exposition
- Discussion closely follows: Piketty and Saez '13

Social Welfare Function

- In general, social planner maximizes $G(v_1, ..., v_n)$
- Social Welfare Functions:
  - Utilitarian: $SWF = \int_n v_n$ or $\sum_n v_n$
  - Rawlsian: $SWF = \min_n(v_1, ..., v_n)$
  - General: $SWF = \int_n G(v_n)$, with $G' > 0$ and $G'' < 0$
  - General Pareto weights: $SWF = \int_n g_n v_n$, with $g_n \geq 0$ exogenously determined
- Social marginal welfare weight: $g_n = G'(v_n)\, u_c^n / \mu$
- The relative value of giving a dollar to person n versus person m:

$$\frac{g_n}{g_m}$$

Revenue Maximization: Laffer Curve

- Use a linear tax τ and demogrant R to maximize revenue [i.e. Rawlsian SWF]
  - Aggregate earnings are: $Z\left(1-\tau, R\left(\tau\right)\right) = \int_n z_n\left(1-\tau, R\left(\tau\right)\right) dF\left(n\right)$
  - Revenue is $R\left(\tau\right) = \tau \cdot Z\left(1-\tau\right)$
- Revenue maximizing rate is:

$$\begin{aligned}\tau^* &= \frac{1}{1+\varepsilon_Z} \\ \text{where } \varepsilon_Z &= \frac{\left(1-\tau\right)}{Z}\frac{\partial Z}{\partial\left(1-\tau\right)}\end{aligned}$$

Optimal Linear Tax Rate

- Government chooses τ to maximize:

$$\int_n G\left[u_n\left((1-\tau)z_n + \tau Z(1-\tau), z_n\right)\right] dF\left(n\right)$$

- Optimal linear tax is:

$$\tau = \frac{1-\bar{g}}{1-\bar{g}+\varepsilon_Z}$$

where $\bar{g} = \int_n \left(z_n/Z\right) g_n dF\left(n\right)$

1. $0 \leq \bar{g} < 1$ if g_n is decreasing with z_n (SMWW falls with consumption).
2. $\bar{g}$ low when (a) inequality is high, (b) $g_n \downarrow$ sharply with c_n
3. Captures the equity-efficiency trade-off robustly ($\tau \downarrow \bar{g}$,$\tau \downarrow \varepsilon$)
4. Rawlsian case: $g_n \equiv 0$ for all $z_n > 0$, so $\bar{g} = 0$ [revenue maximization]
5. Median voter equilibrium $\sim \bar{g} = z_m/Z$

Optimal Top Income Tax Rate

- Now consider the optimal MTR τ for all income above some threshold z^*
- Assume there is a share π^* of individuals earning above z^*
- Let $\bar{z}\left(1-\tau\right)$ be the average earnings above z^*, with elasticity $\bar{\varepsilon} = \left[\left(1-\tau\right)/\bar{z}\right] \cdot d\bar{z}/d\left(1-\tau\right)$
- Note: ε is a mix of income and substitution effects

Optimal Top Income Tax Rate

- At the optimum, top marginal tax rate:

$$\tau = \frac{1 - \bar{g}}{1 - \bar{g} + a \cdot \bar{\varepsilon}}$$

1. Optimal $\tau \downarrow \bar{g}$ [redistributive tastes]
2. Optimal $\tau \downarrow \bar{\varepsilon}$ [efficiency]
3. Optimal $\tau \downarrow a$ [thinness of top tail]
4. Optimal $\tau = 0$ only when $z^* \to z^{Top}$, i.e. $a \to \infty$ [not policy relevant or empirically relevant]
5. Formula robust to heterogeneity, discrete or continuous populations
6. If $\bar{g} \to 0$, top tax rate maximizes revenue [soak the rich]
7. When $z^* = 0$, $a = 1$, and optimal linear tax is obtained

Optimal Top Income Tax Rate

- Empirically: $a = \bar{z}/(\bar{z} - z^*)$ very stable above $z^* = \$400K$, i.e. a Pareto distribution
- Empirically $a \in (1.5, 3)$, US has $a = 1.5$, Denmark has $a = 3$
- Examples:
  - $\bar{\varepsilon} = 0.5$, $\bar{g} = 0.5$, $a = 2 \Longrightarrow \tau^{Top} = 33\%$
  - $\bar{\varepsilon} = 0.5$, $\bar{g} = 0$, $a = 2 \Longrightarrow \tau^{Top} = 50\%$

Optimal Nonlinear Income Tax

- Now consider general problem of setting $T(z)$ [Mirrlees Problem]
- Let $H(z)$ be the income CDF [population normalized to 1] and $h(z)$ its density [endogenous to $T(\cdot)$]
- Let $g(z)$ be the social marginal value of consumption for taxpayers with income z in terms of public funds [formally $g(z) = G'(v_n) \cdot u_c/\mu$]
  - no income effects $\Longrightarrow \int g(z)h(z)dz = 1$
- Redistribution valued $\Longrightarrow g'(z) \leq 0$
- Let $g^+(z)$ be the average social marginal value of c for taxpayers with income above z : $g^+(z) = \int_z^\infty g(s)h(s)\frac{ds}{1-H(z)}$

Optimal Nonlinear Income Tax

- Optimal marginal tax rate at z:

$$T'(z) = \frac{1 - g^{+}(z)}{1 - g^{+}(z) + a(z) \cdot \varepsilon_z}$$

1. Formula does not depend on homogeneity assumption of Mirrlees '71
2. $T'(z) \downarrow \varepsilon_z$ (elasticity efficiency effects) [pure substitution effect]
3. $T'(z) \downarrow a(z) = \frac{zh(z)}{1-H(z)}$ (local Pareto parameter)
4. $T'(z) \downarrow g^{+}(z)$ (redistributive tastes)
5. With no income effects: $g^{+}(z) < 1$ for $z > 0 \rightarrow T'(z) > 0$ [General Mirrlees Result, no EITC]
6. Asymptotics: $g^{+}(z) \rightarrow \bar{g}$, $a(z) \rightarrow a$, $\varepsilon_z \rightarrow \bar{\varepsilon} \Longrightarrow$ Recover top rate formula $\tau = (1 - \bar{g}) / (1 - \bar{g} + a \cdot \varepsilon)$

Extensions

- Income effects can be introduced: higher income effects, all else equal, yield higher tax rates [Saez '01]
- Inverted problem: use current $T(z)$ and $H(z)$ to back out implied $\hat{g}(z)$ [depends on $\hat{\varepsilon}$]
  - Pareto efficient taxation requires $g(z) \geq 0$
- Rent seeking among top earners [Piketty, Saez and Stantcheva '11, Rothschild and Scheuer '11]
- Migration among top earners [Piketty and Saez '13]
- Tax avoidance [Saez, Slemrod and Giertz '12]
- Income Shifting [Piketty and Saez '13]
- Discrete earnings models [Piketty '97 and Saez '02]
- Optimal capital taxation [Saez and Stantcheva '18]

Optimal Transfers: Participation Responses and EITC

- Mirrlees result predicated on assumption that all individuals are at an interior optimum in choice of labor supply
  - Rules out extensive-margin responses
  - But empirical literature shows that participation labor supply responses are important, especially for low incomes
- Diamond (1980), Saez (2002), Laroque (2005) incorporate such extensive labor supply responses into optimal income tax model
- Generate extensive margin by introducing fixed job packages (cannot smoothly choose earnings)

Saez 2002: Participation Model

- Model with discrete earnings outcomes: $z_0 = 0 < z_1 < ... < z_N$
- Tax/transfer T_n when earning z_n, $c_n = z_n - T_n$
- Pure participation choice: skill n individual compares c_n and c_0 when deciding to work
- With participation tax rate τ_n, $c_n - c_0 = z_n \cdot (1 - \tau_n)$
  - Note: $\tau_n = [T_n - T_0]/z_n$
- In aggregate, fraction h_n of population earns z_n, with $\sum_n h_n = 1$
- Participation elasticity is

$$e_n = \frac{(1 - \tau_n)}{h_n} \cdot \frac{\partial h_n}{\partial (1 - \tau_n)}$$

Saez 2002: Participation Model

- Social Welfare function is summarized by social marginal welfare weights at each earnings level g_i
- No income effects $\rightarrow \sum_i g_i h_i = 1 =$ value of public good
- Optimal participation tax:

$$\tau_n = \frac{1 - g_n}{1 - g_n + e_n}$$

  Main result: work subsidies with $T'(z) < 0$ (such as EITC) optimal

  when $g_1 > 1$
- Key requirements in general model with intensive+extensive responses
  - Responses are concentrated primarily along extensive margin
  - Social marginal welfare weight on low skilled workers > 1 (not true with Rawlsian SWF)

Tagging: Akerlof 1978

- We have assumed that $T(z)$ depends only on earnings z
- In reality, govt can observe many other characteristics X also correlated with ability and set $T(z, X)$
  - Ex: gender, race, age, disability, family structure, height,...
- Two major results:
  1. If characteristic X is **immutable** then redistribution across the X groups will be complete [until average social marginal welfare weights are equated across X groups]
  2. If characteristic X can be manipulated but X correlated with ability then taxes will depend on both X and z

Mankiw and Weinzierl 2009

- Tagging with Immutable Characteristics
- Consider a binary immutable tag: **T**all vs. **S**hort
- 1 inch = 2% higher earnings on average (Postlewaite et al. 2004)
- Average social marginal welfare weights $\bar{g}^T < \bar{g}^S$ because tall earn more
- Lump sum transfer from Tall to Short is desirable
- Optimal transfer should be up to the point where $\bar{g}^T = \bar{g}^S$
  - Set optimal non-linear income tax within height groups
- Calibrations show that average tall person ($>$ 6ft) should pay \$4500 more in tax

Problems with Tagging

- Height taxes seem implausible, challenging validity of tagging model
- What is the model missing?
  1. Horizontal Equity concerns impose constraints on feasible policies:
     - Two people earning same amount but of different height should be treated the same way
  2. Height does not *cause* high earnings
     - In practice, tags used only when **causally** related to ability to earn [disability status] or welfare [family structure, # kids, medical expenses]
- Conclude: Mirrlees analysis [$T(z)$] may be most sensible even in an environment with immutable tags

Outline

Taxes and the Economy

Optimal Commodity Taxation

Optimal Income Taxation

First Best Problem

The Social Welfare Function

Limits of the Welfarist Approach

- Welfarism is the dominant approach in optimal taxation
  - Welfarism: social objective is a sole function of individual utilities: $G(u_1, .., u_N)$
- Tractable and coherent framework that captures the equity-efficiency trade-off but generates puzzles:
  1. 100% taxation absent behavioral responses
  2. Whether income is deserved or due to luck is irrelevant
  3. What transfer recipients would have done absent transfers is irrelevant
  4. Tags correlated with ability should be heavily used
- A number of alternatives to welfarism have been proposed
- Saez-Stantcheva '13 (Piketty-Saez '13, section 6 summary) propose a new generalized framework nesting welfarism and many alternatives which can resolve those puzzles

Generalizing the Tax Reform Approach

- Social planner uses generalized social marginal welfare weights $g_n \geq 0$ to value marginal consumption of individual n
  - g_n can vary with $T(z)$ and other economic circumstances
- Optimal tax criterion: $T(z)$ is optimal if:
  - For any budget neutral small tax reform $dT(z)$, $\sum_n g_n dT(z_n) = 0$ with $g_n \geq 0$ generalized social marg. welfare weight on indiv. n

1. Nests welfarist case when $g_n = G_n u_c^n$
2. Generates same optimal tax formulas as welfarist approach
3. Respects (local) constrained Pareto efficiency ($g_n \geq 0$)
4. No social objective is maximized [Instead local tax reforms considered]

Application 1: Optimal Tax with Fixed Incomes

- Utilitarian approach has degenerate solution with 100% taxation when $u''(c) < 0$
  - Public may not support confiscatory taxation even absent behavioral responses
- Generalized social marginal welfare weights: $g_n = g(c_n, T_n)$
  - $g_c(c, T) < 0$ (ability to pay)
  - $g_T(c, T) > 0$ (contribution to society)
- Optimum: $g(z - T(z), T(z))$ equalized across z:

$$T'(z) = \frac{1}{1 - g_T / g_c}$$

and $0 \leq T'(z) \leq 1$

Application 1: Optimal Tax with Fixed Incomes

- Preferences for redistributions embodied in $g(c, T)$
- Polar cases:
  1. Utilitarian case: $g(c, T) = u\prime(c) \downarrow c \Longrightarrow T'(z) \equiv 1$
  2. Libertarian case: $g(c, T) = g(T) \uparrow T \Rightarrow T'(z) \equiv 0$
- SS '13 use Amazon mTurk online survey to estimate $g(c, T)$
- They find that revealed preferences depend on both c and T :
  - $\{z = \$40K, T = \$10K, c = \$30K\}$ more deserving than $\{z = \$50K, T = \$10K, c = \$40K\}$
  - $\{z = \$50K, T = \$15K, c = \$35K\}$ more deserving than $\{z = \$40K, T = \$5K, c = \$35K\}$

Application 2: Deserved vs. Luck Income

- Taxing luck income (Paris Hilton) is fair while taxing deserved income (Steve Jobs) is not
- Suppose $z = w + y$ with w deserved income and y luck income (w,y mix not observable)
- Person is deserving if:
  - $c = z - T \leq w + \mathbb{E}[y]$ with $\mathbb{E}[y]$ average luck income
  - $\implies g_n = 1$ if $c_i \leq w_i + \mathbb{E}[y]$
  - $g_n = 0$ if not
- $\Pr[g_n = 1 | w + y = z]$ provides micro-foundation for $g(c, T)$ increasing in T
- Beliefs in share of income due to luck at each income level is key

Application 3: "Free Loaders"

- SS '13 online survey shows strong public preference for redistributing toward deserving poor (unable to work or trying hard to work) rather than undeserving poor (who would work absent transfers)
- Generalized social welfare weights can capture this by setting $g_n = 0$ on free loaders (i.e. transfer recipients who would have worked absent the transfer)
  1. Behavioral responses reduce desirability of transfers (over and above standard budgetary effect)
  2. In-work benefit – $T'(0) = (g_0 - 1)/(g_0 - 1 + e_0) < 0$ at bottom – becomes optimal in Mirrlees (1971) optimal tax model if $g_0 < 1$

Link with other Social Justice Principles

- Various alternatives to welfarism have been proposed
- Each alternative can be recast in terms of implied generalized social marginal welfare weights (as long as it generates constrained Pareto efficient optima)
- In all cases, we can use simple and tractable optimal income tax formula for heterogeneous population from Saez Restud'01 (case with no income effects):

$$T'(z) = \frac{1 - G(z)}{1 - G(z) + \alpha(z) \cdot e}$$

  with $G(z)$ average of g_n above z
- g_n average to one in the full population and hence $G(0) = 1$

Link with other Social Justice Principles

1. Rawlsian: g_n concentrated on worst-off individual $\Longrightarrow G(z) = 0$ for $z > 0$ and $T'(z) = 1/(1 + \alpha(z) \cdot e)$ revenue maximizing
2. Libertarian: $g_n \equiv 1 \Longrightarrow G(z) \equiv 1$ and $T'(z) \equiv 0$
3. Equality of Opportunity: (Roemer '98) g_n concentrated on those coming from disadvantaged background. $G(z)$: relative fraction of individuals above z coming from disadvantaged background
   - $G'(z) < 0$ for reasons unrelated to diminishing marginal utility