Econ 230B
Spring 2017
Emmanuel Saez
Yotam Shem-Tov, shemtov@berkeley.edu

Problem Set 1

1. Lorenz Curve and Gini Coefficient

a) See Figure at the end.

b+c) The results using the Pareto interpolation are:

	2004	2014
Top 10% share	0.44	0.47
Top 1% share	0.19	0.20
90/50 ratio	3.29	3.17

Elaborate explanation on Pareto interpolations

We will use a Pareto interpolation following the method discussed in section 3 in Atkinson et al. (2011) and section 5.1 in Atkinson (2005). We will begin by describing the Pareto distribution and some of its useful properties, and then discuss the Pareto interpolation procedure.
The Pareto cumulative distribution function $F(y)$ for y is,

$$F(y) = 1 - \left(\frac{k}{y}\right)^{\alpha}, \quad \alpha > 1,\ k > 0$$
$$\Rightarrow f(y) = \frac{\alpha k^{\alpha}}{y^{1+\alpha}}$$

where k and α are given parameters, α is referred to as the Pareto parameter. The α parameter determines the thickness of the tail of the distribution.
The key property of the Pareto distribution is that the ratio of average income $y^*(y)$ of individuals above y to y does not depend on the income threshold of y (k)[1],

$$y^*(y) = \frac{\int_{z>y} z f(z)dz}{\int_{z>y} f(z)dz} = \frac{\int_{z>y} z\alpha k^{\alpha}/z^{1+\alpha}dz}{\int_{z>y} \alpha k^{\alpha}/z^{1+\alpha}dz} = \frac{\int_{z>y} 1/z^{\alpha}dz}{\int_{z>y} 1/z^{1+\alpha}dz} \tag{1}$$
$$= \frac{\frac{1}{-(\alpha-1)z^{\alpha-1}}\Big|_y^{\infty}}{\frac{1}{-\alpha z^{\alpha}}\Big|_y^{\infty}} = \frac{0 - \frac{1}{(\alpha-1)y^{\alpha-1}}}{0 - \frac{1}{\alpha y^{\alpha}}} = y \cdot \alpha/(\alpha-1)$$
$$\Rightarrow \frac{y^*(y)}{y} = \beta, \quad with \quad \beta = \alpha/(\alpha-1)$$

[1] Using the notation from Atkinson et al. (2011)

The parameter β has a more intuitive interpretation than α, if $\beta = 2$ the average income of individuals with income above $\$100,000$ is $\$200,000$. A higher β means a flatter upper tail of the income distribution. We will use Atkinson (2005) notation. Denote by $H(y)$ the share of individual with income higher than y,

$$\begin{aligned} H(y) = 1 - F(y) &= \left(\frac{k}{y}\right)^{\alpha} \\ &= Ay^{-\alpha}, \quad with \quad A = k^{\alpha} \end{aligned}$$

Denote by $G(y)$ the total income of received by individuals with income y or higher. According to the properties of the Pareto distribution,

$$G(y) = \frac{\alpha}{\alpha - 1} Ay^{-(\alpha-1)}$$

The share in total income of people with incomes y or higher can be obtained by dividing $G(y)$ by μ, where μ is the mean income, denote by $\Omega(y) = \frac{G(y)}{\mu}$.

Let y_p the desired income percentile. There are to kinds of interpolations:

1. When we observe tabulations of incomes shares lower and higher then y_p, in this case we can use interpolation between the two adjacent tabulations that y_p is in the middle of them. For example we observe the total income of the top 3.677% and the top 0.7688%, however we do not observe the total income of the top 1%.

2. The second scenario is when we want to estimate the share in total income of a percentile we observe only percentiles that are higher than it. For example the smallest percentile we observe is 0.0123%, hence if we want to estimate the share in total income of the top 0.001% we will have only tabulations from one side. This is referred as open range tabulation. In this exercise all the tabulations we are asked to perform are from the first type.

Pareto interpolations

I will begin by interpolating the share in total income of the top 1%. Denote by G_1, G_2, and G^* the share in total of the top $0.768\%, 3.677\%$ and 1% respectively. Denote by H_1 and H_2 the $H(y)$ of the two tabulations to the left and right of $H(y) = 1$ (e.g., $H_1 = 0.00768$ and $H_2 = 0.03677$).

It follows from the Parto distribution that [2]

[2]Proof: Let y_1 and y_2 be two income levels such that $y_1 > y_2$ and therefore $H(y_1) < H(y_2)$.

$$\begin{aligned} \frac{\log(H_1/H_2)}{\log(G_1/G_2)} &= \frac{\log\left(Ay_1^{-\alpha}/Ay_2^{-\alpha}\right)}{\log\left(\frac{\frac{\alpha}{\alpha-1}Ay_1^{-(\alpha-1)}}{\frac{\alpha}{\alpha-1}Ay_2^{-(\alpha-1)}}\right)} = \frac{\log\left((y_1/y_2)^{-\alpha}\right)}{\log\left((y_1/y_2)^{-(\alpha-1)}\right)} \\ &= \frac{-\alpha\log\left((y_1/y_2)\right)}{-(\alpha-1)\log\left((y_1/y_2)\right)} = \frac{\alpha}{\alpha-1} \quad \square \end{aligned}$$

$$\frac{\alpha}{\alpha-1}=\frac{\log\left(H_1/H_2\right)}{\log\left(G_1/G_2\right)} \tag{2}$$

Therefore using equation (2) we can estimate $\frac{\alpha}{\alpha-1}$ for the top 1 percent. Note, that this is making a Pareto interpolation between the two tabulations closest to the desired quantity.

$$\beta=\frac{\log\left(H_1/H_2\right)}{\log\left(G_1/G_2\right)}$$
$$\Rightarrow \alpha=\frac{\beta}{\beta-1}$$

Next we use G_1 and H_1 to approximate the total earnings of the top 1% (it does not matter if we choose H_2 and G_2, the result would have been the same).

$$\beta \approx \frac{\log\left(H_1/0.01\right)}{\log\left(G_2/G^*\right)} \Rightarrow G^*=\log\left(G_2/G^*\right)=\frac{1}{\beta}\cdot\log\left(H_1/0.01\right)$$
$$\Rightarrow G_2/G^*=\exp\left(\frac{1}{\beta}\cdot\log\left(H_1/0.01\right)\right)=\left(H_1/0.01\right)^{\frac{1}{\beta}}$$
$$G^*=G_2\cdot\left(H_1/0.01\right)^{\frac{1}{\beta}}$$

Next we use the same procedure as before to estimate α for each percentile, and then use the CDF of the Pareto distribution. Denote by p the percentile of interest and by y_p income of the individual at the p percentile. Given a tabulation we can calculate α. Next we extrapolate the parameter k (or A in Atkinson's notation) for each of the tabulations above and below the desired p,

$$k=(1-F(y))^{1/\alpha}\cdot y$$

After we have k and α the desired income threshold, y_p is,

$$y_p=\frac{k}{(1-F(y_p))^{1/\alpha}} \tag{3}$$

We need to decide which of the tabulations to use in order to calculate k. Pikkety 2001, Appendix B, (Page 598 in the book) suggest to choose the tabulation that is closest to the desired percentile. It is clear that the choice of which tabulation to use has a significant impact on the results.

d) See second figure at the end of the solutions for the Lorenz curve.

e) Mis-ranking always under-estimates the Gini coefficient (as the Lorenz curve is furthest away from the diagonal precisely when individuals are perfectly ranked by income).

2. Optimal income tax

An economy is populated by individuals with preferences over consumption and labor. They have utility $u_i(c, y)$ where y is income, $u_c(c, y) > 0$ and $u_y(c, y) < 0$. Suppose the tax schedule in place has a constant marginal tax rate τ above a fixed threshold y^*. The government wants to choose τ to maximize the tax revenue raised from top earners.

(a) As we saw in class, the tax rate that maximizes revenues depends on a Pareto parameter a and the elasticity of total income of the top earners who are in the top bracket, ε. Provide intuition about why ε is a mix of substitution and income effects.

Increasing τ by $d\tau$ generates a negative substitution effect (less slope) leading to less work, and a negative income effect leading to more work. Hence, ε is a mix of substitution and income effects.

(b) The individual solves the following utility maximization problem:

$$\max_{c,y} u_i(c, y)$$

subject to:

$$c = (1 - \tau)y + I$$

Denote by $y_i(1 - \tau, I)$ the Marshallian income supply. The uncompensated elasticity of labor supply with respect to $1 - \tau$ is $\varepsilon_i^u = \big(\partial y_i/\partial(1 - \tau)\big)\big((1 - \tau)/y_i\big)$. We denote by $\eta_i = (1 - \tau)\partial y_i/\partial I$ the income parameter.
Suppose a government advisor suggests to run an experiment where the top tax rate τ (above y^*) is raised by $d\tau$. The advisor claims that the response dy_i can be rewritten as a function of ε_i^u and η_i. Is the advisor right? If yes, show how dy_i depends on ε_i^u and η_i

Let $z_i(1, I)$ be the earnings supply function obtained from solving the individual utility maximization problem under a linear tax:

$$\max_{c,y} u_i(c, y)$$

subject to:

$$c = (1 - \tau)y + I$$

We denote by ε_i^u the uncompensated elasticity of y_i with respect to $(1 - \tau)$ and by $\eta_i = (1 - \tau)\partial y_i/\partial I$ the income effect parameter. As a simple graph shows (see Saez Restud01 Section 3), the reform changes $(1 - \tau)$ by $-d\tau$ and changes I by $dI = y^* d\tau$. Note, that the virtual income is defined as $Iy^*\tau$ (in Saez 2001 it is written as $I = \hat{y}\tau$), and as we assume that y^* is fixed it follows immediately that $dI = y^* d\tau$. Hence, we have:

$$dy_i = -d\tau \frac{\partial y_i}{\partial(1-\tau)} + y^* d\tau \frac{\partial y_i}{\partial I} = -\frac{d\tau}{1-\tau} y_i \frac{1-\tau}{y_i} \frac{\partial y_i}{\partial(1-\tau)} + \frac{d\tau}{1-\tau} y^*(1-\tau) \frac{\partial y_i}{\partial I} = -\frac{d\tau}{1-\tau} y^i \varepsilon_i^u + \frac{d\tau}{1-\tau} y^* \eta_i$$

(c) Using the expression derived in point b) write ε as a function of the Pareto parameter $a = y^m/(y^m - y^*)$ and a weighted average of the uncompensated elasticities and income effect parameters. Why are uncompensated elasticities weighted by incomes y_i, while the η_is are not?

Recall that the elasticity ε is defined as

$$\varepsilon = \frac{1-\tau}{\sum_{y_i \geq y^*} y_i} \frac{\sum_{y_i \geq y^*} dy_i}{d(1-\tau)}$$

Hence we have

$$\varepsilon = \frac{1-\tau}{\sum_{y_i \geq y^*} y_i} \sum_{y_i \geq y^*} \left[\frac{1}{1-\tau} y_i \varepsilon_i^u - \frac{1}{1-\tau} y^* \eta_i\right] = \frac{\sum_{y_i \geq y^*} y_i \varepsilon_i^u}{\sum_{y_i \geq y^*} y_i} - \frac{z^*}{z^m} \frac{\sum_{y_i \geq y^*} \eta_i}{N}$$

with N number of top bracket taxpayers. Hence $\varepsilon = \hat{\varepsilon}^u - \frac{a-1}{a}\hat{\eta}$ the income weighted average of ε_i^u and $\hat{\eta}$ the straight average of η_i among top bracket taxpayers.
The uncompensated elasticity is income weighted because those with higher income should count more in the response. In contrast, the income effect parameter is not an elasticity and hence should not be income weighted.

(d) Now suppose the utility is logarithmic in consumption and exponential in income. It takes the following form:

$$u_i(c, y) = \log c - \phi_i y^{1+\frac{1}{\varepsilon}}$$

where ϕ_i can vary across individuals and captures heterogeneity in the disutility from labor. Derive the uncompensated elasticity, income, and compensated elasticity parameters (i.e., ε_i^u, η_i, ε_i^c) by solving the utility maximization problem of the individual under the linear tax and the same budget constraint as above.

Suppose each individual utility takes the form $u_i(c, y) = \log c - \phi_i y^{1+\frac{1}{\varepsilon}}$ where phi_i is a parameter (that can vary across individuals). The individual solves

$$\max_{c,y} \log((1-\tau)y + I) - \phi_i y^{1+\frac{1}{\varepsilon}}$$

The individual FOC in y is:

$$\frac{1-\tau}{(1-\tau)y+I} = \phi_i(1+\frac{1}{\varepsilon})y^{\frac{1}{\varepsilon}}$$

or

$$\log(1-\tau) - \log((1-\tau)y+I)) = \frac{1}{\varepsilon}\log(y) + \log(\phi_i(1+\frac{1}{\varepsilon}))$$

We now derive elasticities at the limit where $y \to \infty$ since we know from the FOC that when $\phi \to 0$ the agent supplies and infinite amount of labor. This defines implicitly $y^i(1-\tau, I)$ with the following comparative statics. A small change dI leads to dz_i such that:

$$dy_i\Big[\frac{1}{\varepsilon y} + \frac{1-\tau}{(1-\tau)y+I}\Big] = -\frac{dI}{(1-\tau)y+I}$$

Hence

$$\eta_i = (1-\tau)\frac{dy_i}{dI} = -\frac{1}{1+\frac{y+I/(1-\tau)}{\varepsilon y}} \to_{y\to\infty} -\frac{\varepsilon}{\varepsilon+1}$$

A small change $d(1-\tau)$ leads to dy_i such that:

$$dy_i\Big[\frac{1}{\varepsilon y} + \frac{1-\tau}{(1-\tau)y+I}\Big] = \frac{d(1-\tau)}{1-\tau} - \frac{yd(1-\tau)}{(1-\tau)y+I}$$

Hence

$$(1-\tau)\frac{dz_i}{d(1-\tau)}\Big[\frac{1}{\varepsilon y} + \frac{1}{(1-\tau)y+I}\Big] = 1 - \frac{y}{(1-\tau)y+I}$$

$$\varepsilon_i^u = \frac{1-\tau}{y}\frac{dz_i}{d(1-\tau)} = -\frac{1-\frac{z}{(1-\tau)y+I}}{\frac{1}{\varepsilon}+\frac{y}{(1-\tau)y+I}} \to_{y\to\infty} 0$$

(e) Study what happens to η_i and ε_i^u when ϕ_i becomes small (find their limits). Using the relation found previously, write the optimal top tax rate formula as a function of ε and the Pareto parameter a when ϕ_i is small.
The parameter ε is the Frisch elasticity of labor supply for this class of utility functions. Suppose we calibrate the parameters a and ε such that $a = 1.5$ and $\varepsilon = 1$. What is the optimal τ? What is the optimal τ when ε is very large? Discuss why the optimal tax rate is high even with a large Frisch elasticity.

Use c) to obtain an optimal top tax rate formula as a function of ε and the Pareto parameter a in that case. We have $\varepsilon = \hat{\varepsilon}_u - \frac{a-1}{a}\hat{\eta} \simeq \frac{a-1}{a}\frac{\varepsilon}{\varepsilon+1}$ and hence:

$$\tau = \frac{1}{1+a\varepsilon} = \frac{1}{1+(a-1)\frac{\varepsilon}{1+\varepsilon}} = \frac{\varepsilon+1}{1+a\varepsilon}$$

If $a = 1.5$ and $\varepsilon = 1$, we have $\tau = 1/(1+0.5\times 0.5) = 80\%$.
With $\varepsilon = \infty$, we get $\tau = 1/(1+0.5) = 66.6\%$. This utility specification has zero uncompensated elasticity and hence large income effects when the compensated elasticity is large. As a result, the tax rate is substantial, even with a very large Frisch elasticity.

3. Optimal linear income tax

Suppose that utility is quasi-linear and takes the form: $u(c,l) = c - \frac{l^{1+\epsilon}}{1+\epsilon}$ with $\epsilon > 0$. Each individual earns income $y = wl$ and consumes $c = y - T(y)$. The wage rate w can be interpreted as a measure of skills and is distributed with density $f(w) > 0$ over $[0,\infty)$. The total population is normalized to one so that $\int_0^\infty f(w)dw = 1$

(a) Suppose the tax schedule is linear with a flat tax rate τ. The tax is hence $T(y) = -S + \tau y$ where $S > 0$ is the transfer that the individual receives when labor supply is zero ($T(0) = -S$). Find the optimal labor supply choice as a function of the parameters S and $w(1-\tau)$. Also, derive the uncompensated and compensated elasticities of labor supply as a function of ϵ and find the income effect parameter.

$\max_l wl(1-\tau) + S - l^{1+\epsilon}/(1+\epsilon) \Longrightarrow w(1-\tau) = l^\epsilon \Longrightarrow l = \left[w(1-\tau)\right]^{1/\epsilon}$ so $\varepsilon^u = \varepsilon^c = 1/\epsilon$ and $\eta = 0$. Let us denote by $\varepsilon = 1/\epsilon$ the common compensated and uncompensated elasticity.

(b) Assume that taxes are entirely rebated to the individuals in the economy. We have that $S = \tau Y$, where Y is average earnings in the economy. Find the optimal tax rate τ in the case where the government only cares about the worst-off individual (i.e. the government is Rawlsian) and in the case where the government maximizes the sum of utilities (i.e. the government is utilitarian). Always explain the intuition behind your results.

$$\max_\tau wl(1-\tau)^{\frac{1}{\epsilon}} \int w^{1+\frac{1}{\epsilon}} f(w)dw \Longrightarrow \tau^* = \frac{\epsilon}{\epsilon+1} = \frac{1}{1+\varepsilon}$$

Worst off individual has $w = 0$ and hence $l = 0$ and utility $u = S = \tau Y$ so Rawlsian optimal rate maximizes tax revenue (to maximize S) and is set at $\tau^* = \frac{\epsilon}{\epsilon+1}$ from (b). Given that all utilities are linear, there is no concern for redistribution and hence the optimal utilitarian tax rate is zero.

(c) Do points (a)-(b) again using utility function $u(c,l) = \log(c) - l$. If exact analytical expressions are not possible to derive, just provide implicit formulas with economic explanation. Is this utility function more or less realistic than the one used in questions (a)-(b)?

Go back to utility function $u(c,l) = c - \frac{l^{1+\epsilon}}{1+\epsilon}$. We now study an economy with two tax brakets such that:

$$T(y) = \begin{cases} -S + \tau_1 y & \text{if } y \leq \hat{y} \\ -S + \tau_1 \hat{y} + \tau_2 (y - \hat{y}) & \text{if } y > \hat{y} \end{cases}$$

$-S$ is the transfer to non-working individuals.

With utility $\log(c) - l$, we have $\max_l \log\big(wl(1-\tau)+S\big) - l \Longrightarrow w(1-\tau)/\big[w(1-\tau)l + S\big] = 1$ so that $l = 1 - S/\big[w(1-\tau)\big]$. Note that $l = 0$ when $w(1-\tau) \leq S$. Income effect $\eta = -1$, $\varepsilon^u = S/\big[S - w(1-\tau)\big]$, $\varepsilon^c = \varepsilon^u - \eta = w(1-\tau)/\big[S - w(1-\tau)\big]$. $\tau^* = 1/(1+\hat{\varepsilon}^u)$ where $\hat{\varepsilon}^u$ is the (income weighted) average uncompensated elasticity. $\hat{\varepsilon}^u$ does not have a simple analytic expression. Worst-off individual has $l = 0$ and utility $u = \log(S)$ so government wants to maximize S which is done by maximizing tax revenue with $\tau^* = 1/(1+\hat{\varepsilon}^u)$. In utilitarian case, the optimal τ is given by $\tau = (1-\hat{g})/(1-\hat{g}+\varepsilon)$ as seen in class notes with ε a mix of uncompensated and income effects (see Piketty-Saez handbook chapter for details). There is no simple analytic expression.

(d) Plot the budget constraint on a graph with axes (l, c).

(e) Suppose that $\tau_1 < \tau_2$. Find the optimal labor supply and earnings for an individual with wage w. Consider the three cases where the individual is in the bottom bracket, the top bracket, or exactly at $\hat{y}$.

The budget has a kink generating bunching at $\hat{y}$.
Case 1 (first bracket): $w \leq \underline{w}$: $l = \big[w(1-\tau_1)\big]^{\varepsilon}$ with $\underline{w}$ s.t. $\underline{w}^{1+\varepsilon}(1-\tau_1)^{\varepsilon} = \hat{y}$
Case 2 (bunching at $\hat{y}$): $\underline{w} \leq w \leq \hat{w}$: $l = \hat{y}/w$
Case 3 (top bracket): $w \leq \overline{w}$: $l = \big[w(1-\tau_2)\big]^{\varepsilon}$ with $\overline{w}$ s.t. $\overline{w}^{1+\varepsilon}(1-\tau_2)^{\varepsilon} = \hat{y}$

Suppose that there are 3 types of individuals: disabled individuals unable to work $w_0 = 0$, low skilled individuals with wage rate w_1, and skilled individuals with wage rate w_2. We assume that $w_1 < w_2$. The fractions of disabled, low skilled, and high skilled in the population are respectively λ_0, λ_1 and λ_2 such that $\lambda_0 + \lambda_1 + \lambda_2 = 1$. Further assume that low skilled workers are always in the bottom bracket and that high skilled workers are always in the top bracket.

(f) Find the tax rate τ_2^* that maximizes taxes collected from the high skilled, assuming that S, τ_1, and $\hat{y}$ are given. Express it as a function of ϵ and $\hat{y}$.

Amount of bunching is proportional to ε (see Saez AEJ:EP10 for details):

$$\frac{\overline{w}^{1+\varepsilon}}{\underline{w}^{1+\varepsilon}} = \left(\frac{1-\tau_1}{1-\tau_2}\right)^{\varepsilon}$$

$$T = \lambda_1 \tau_1 (1-\tau_1)^\varepsilon w_1^{1+\varepsilon} + \lambda_2 \Big\{ \tau_1 \hat{y} + \tau_2 \Big[(1-\tau_2)^\varepsilon w_2^{1+\varepsilon} - \hat{y} \Big] \Big\}$$

Take the FOC wrt to τ_2 to get:

$$\tau_2^* / (1-\tau_2^*) = (1/\varepsilon) \big[y_2 - \overline{y} \big] / y_2$$

(g) Compute the tax rate τ_1 that maximizes total taxes collected taking S and $\hat{y}$ as given and setting $\tau_2 = \tau_2^*$ (the optimal tax rate you found in the previous question). Explain why (intuitively) $\tau_2^* < \tau^* < \tau_1^*$, where τ^* is the one computed in question (b).

Take the FOC wrt to τ_1 to get:

$$\tau_1^* / (1-\tau_1^*) = (1/\varepsilon) \big[1 + \lambda_2 \overline{y} / (\lambda_1 y_1) \big]$$

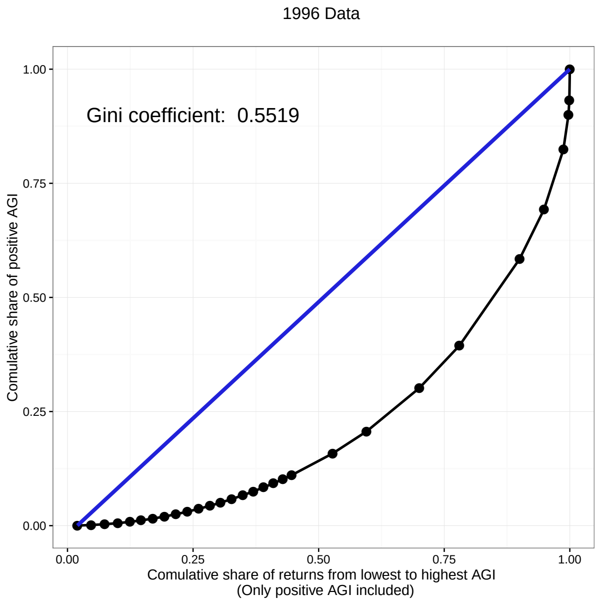
In order to explain $\tau_2^* < \tau^* < \tau_1^*$:

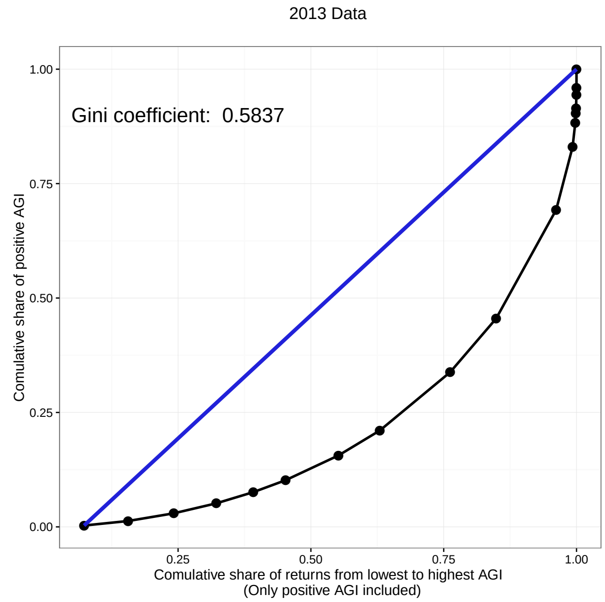
1. Increasing the flat tax rate τ creates a mechanical increase in revenue proportional to average earnings and creates a negative behavioral response proportional to average earnings as well.

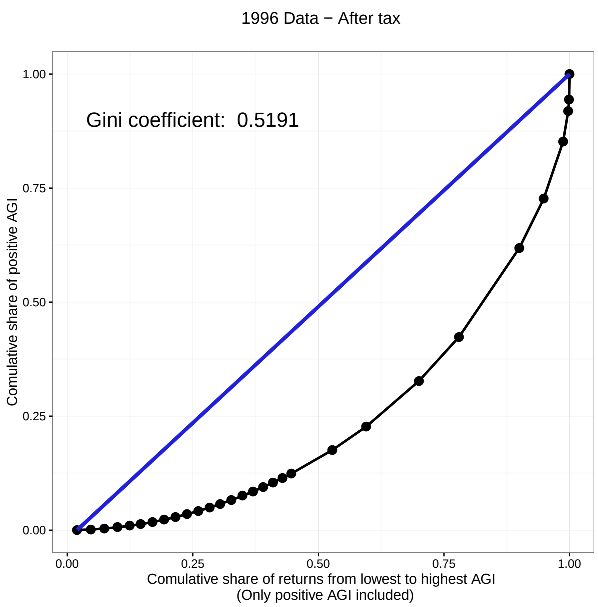
2. Increasing the tax rate τ_2 in the top bracket creates a mechanical increase in revenue proportional to $(y_2 - \overline{y})$ but creates a negative behavioral response proportional to y_2.

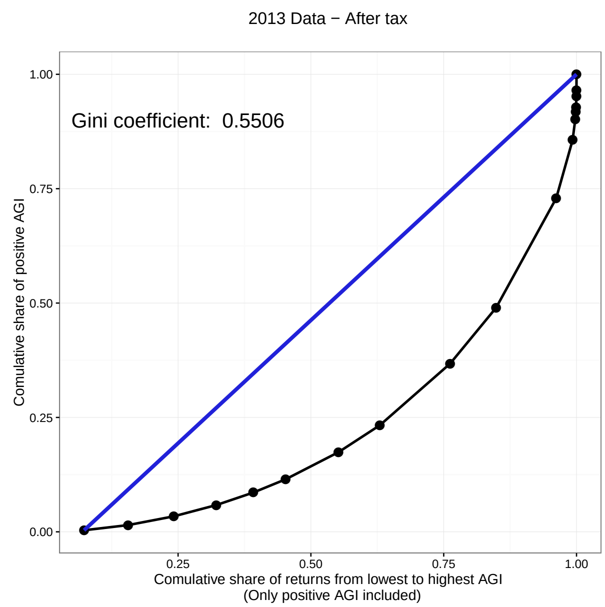
3. Increasing the tax rate τ_1 in the bottom bracket creates a mechanical increase in revenue proportional to y_1 and creates a negative behavioral response proportional to y_1. However, the tax rate increase also raises more tax from high skilled worker with no negative behavioral response (inframarginal tax).

Figures


1996 Data
Gini coefficient: 0.5519
Comulative share of positive AGI
Comulative share of returns from lowest to highest AGI
(Only positive AGI included)
0.00
0.25
0.50
0.75
1.00


2013 Data
Gini coefficient: 0.5837
Comulative share of positive AGI
Comulative share of returns from lowest to highest AGI
(Only positive AGI included)
0.00
0.25
0.50
0.75
1.00


1996 Data − After tax
Gini coefficient: 0.5191
Comulative share of positive AGI
Comulative share of returns from lowest to highest AGI
(Only positive AGI included)
0.00
0.25
0.50
0.75
1.00


2013 Data − After tax
Gini coefficient: 0.5506
Comulative share of positive AGI
Comulative share of returns from lowest to highest AGI
(Only positive AGI included)
0.00
0.25
0.50
0.75
1.00

References

Atkinson, Anthony, "Top Incomes in the UK over the 20th Century," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 2005, *168* (2), 325–343.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez, "Top Incomes in the Long Run of History," *Journal of Economic Literature*, 2011, *49* (1), 3–71.