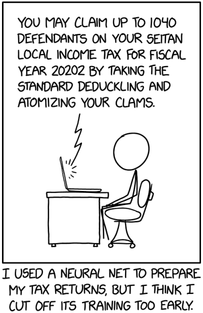
Income Taxes

Kyle Coombs (adapted from many sources)

Vassar College

October 2, 2025



Why do we tax income? How should we?

- Who should pay more in taxes?
- What is a progressive tax?
- How does income taxation increase equity?
- Imagine we tax 100% of income above $1 million. What would happen?

Lecture goals

- Characterize the equity and efficiency tradeoffs in optimal income taxation
- Separate the normative and positive components of the sufficient statistics approach to optimal income taxation
- Learn to evaluate top marginal tax rates

Labor Income Taxes

Labor Income Taxes: payroll and income taxes tax labor

- Labor income taxes reduce the marginal returns to work
- **Labor Supply Elasticity:** (ε^s) how do your hours worked respond to changes in your post-tax wage

$$\varepsilon^s = \frac{\partial \text{Hours Worked}}{\partial \text{Post-Tax Wage}} \times \frac{\text{Post-Tax Wage}}{\text{Hours Worked}} \tag{1}$$

Is $\varepsilon > 0$ or < 0?

Labor Income Taxes

Labor Income Taxes: payroll and income taxes tax labor

- Labor income taxes reduce the marginal returns to work
- **Labor Supply Elasticity:** (ε^s) how do your hours worked respond to changes in your post-tax wage

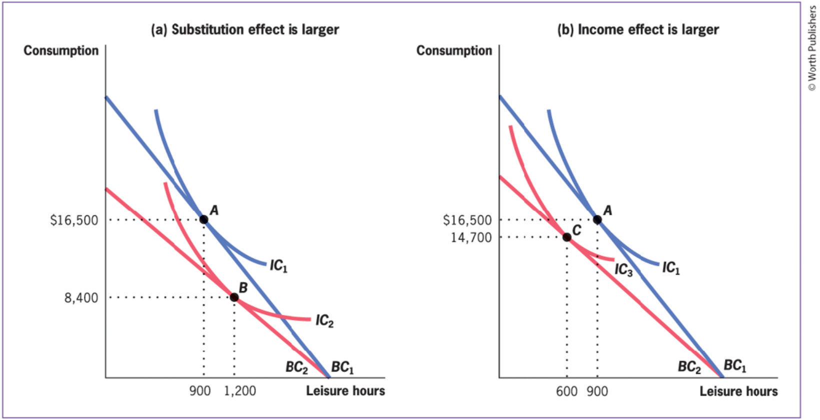
$$\varepsilon^s = \frac{\partial \text{Hours Worked}}{\partial \text{Post-Tax Wage}} \times \frac{\text{Post-Tax Wage}}{\text{Hours Worked}} \tag{1}$$

Is $\varepsilon > 0$ or < 0? It depends!

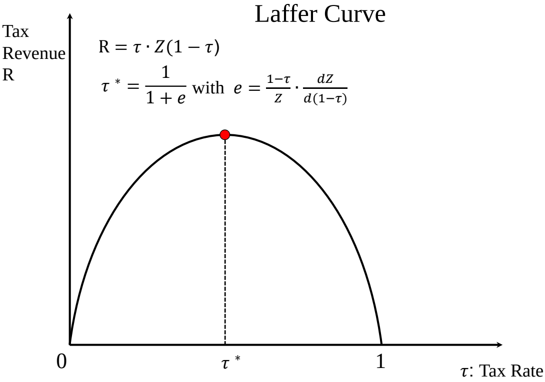
- **Substition effect:** Taxes decrease the price of leisure, so people work less.
- **Income effect:** Taxes reduce income, so people work more if leisure is a normal good.

Effects are opposite-signed, so the theoretical impact of taxation on labor supply is ambiguous.

Substitution and Income Effects on Labor Supply



- (a) When substitution effect is larger: ↓ hours worked
- (b) When income effect is larger: ↑ hours worked



The Laffer curve shows the relationship between tax rates and tax revenue, where e is taxable income elasticity. For low tax rates, tax revenue increases with tax rate. For high tax rates, tax revenue decreases with tax rate.

Optimal income tax: challenge

- Many people have unobserved ability to pay
  - Horizontal equity: Equal ability, equal tax rate
  - Vertical equity: Higher ability, higher tax rate
- Gov't observes income $z = w \cdot L$, where L is labor supply
- z doesn't reveal ability; same z can be high w, low L, or vice versa
- Government can tax income; prefers redistribution
- High taxes on high ability individuals may reduce their work effort, making them appear low ability
- *What is the optimal tax policy* (i.e., tax function $T(z)$)?
- Complex: Vickrey (1945, Nobel), Mirrlees (1971, Nobel)

Optimal income tax goals

- Income taxation: tradeoff between equity and efficiency
- Reminder: Pareto efficiency says nothing about equity
- Given a social welfare function, optimal income tax theory delivers the best tradeoff between equity and efficiency
- Most common SWF: Utilitarian (sum of individual utilities)

Motivating example[1]

- Two types of people, high and low-income (H and L) each with the same utility that exhibits diminishing marginal utility
- Each receives z_i: income net of taxes/transfers
- Gov't taxes H; transfers to L. Has a utilitarian SWF
- Labor demand for H is perfectly elastic (tax burden on H)
- Change in welfare from a tax is captured by:

$$\sum_{i \in \{L,H\}} \left[(\Delta z_i)(MU_i) \right] \Rightarrow MU_L \times \Delta z_L - MU_H \times \Delta z_H \qquad (2)$$

MU_i is the marginal utility of income for person i

[1] Adapted from Jon Bakija and Emmanuel Saez

Motivating example[1]

- Two types of people, high and low-income (H and L) each with the same utility that exhibits diminishing marginal utility
- Each receives z_i: income net of taxes/transfers
- Gov't taxes H; transfers to L. Has a utilitarian SWF
- Labor demand for H is perfectly elastic (tax burden on H)
- Change in welfare from a tax is captured by:

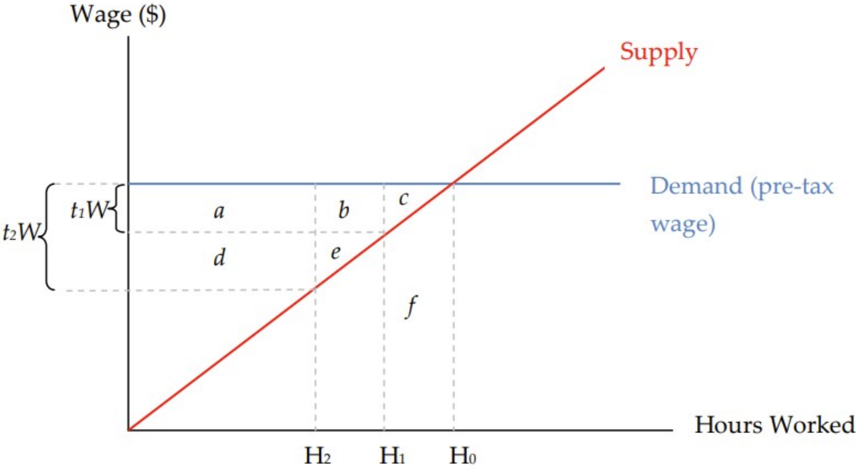
$$\sum_{i \in \{L,H\}} [(\Delta z_i)(MU_i)] \Rightarrow MU_L \times \Delta z_L - MU_H \times \Delta z_H \qquad (2)$$

MU_i is the marginal utility of income for person i

- If eq 2 is positive, the tax is welfare improving (and vice versa)

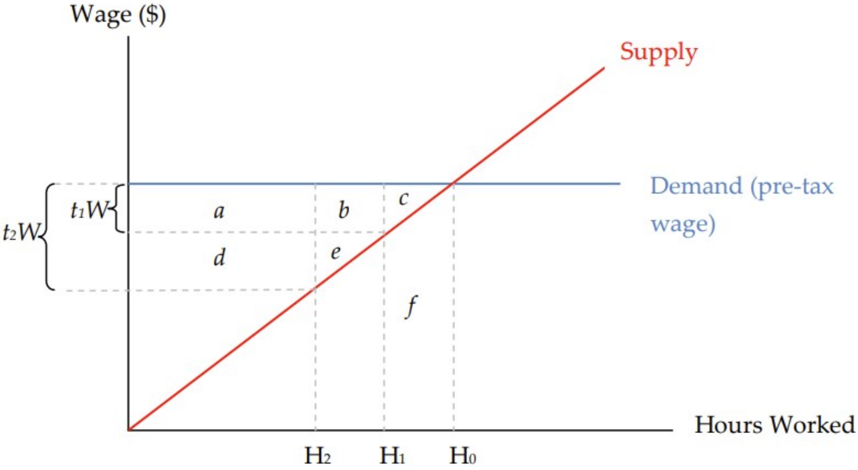
[1] Adapted from Jon Bakija and Emmanuel Saez

Behavioral responses with upward sloping labor supply



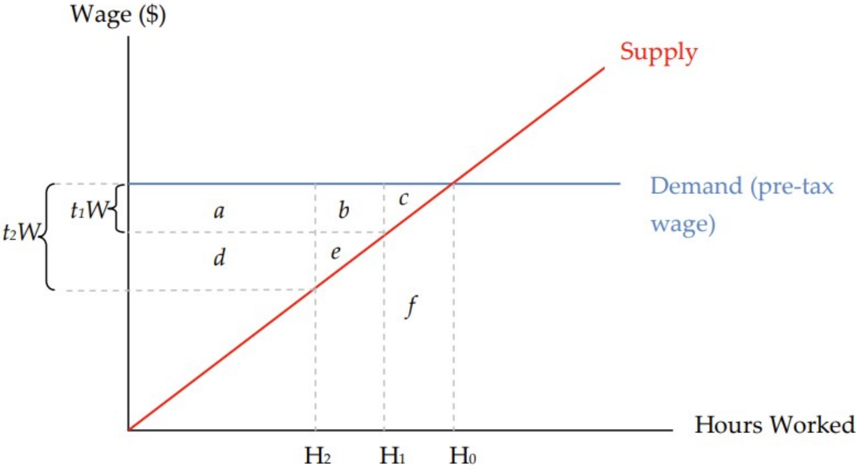
Tax of t_1 raises $a + b$ given to L, H loses $a + b + c$

Behavioral responses with upward sloping labor supply



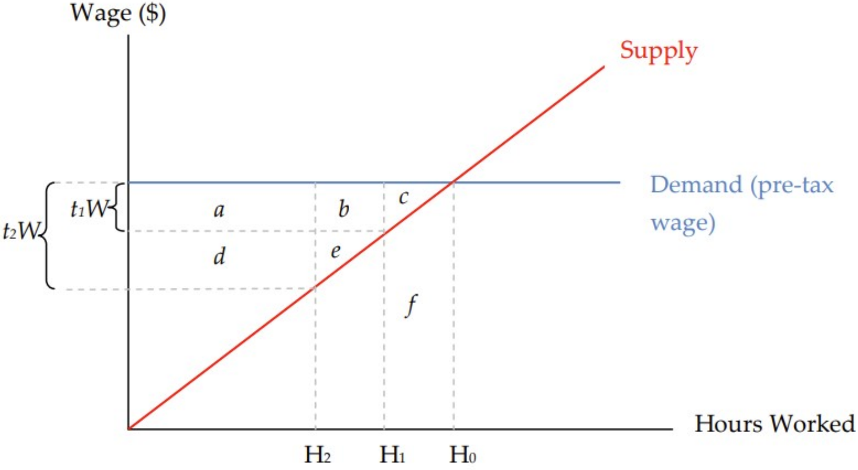
If $MU_L^* \cdot (a+b) > MU_H^* \cdot (a+b+c)$; tax is welfare improving

Behavioral responses with upward sloping labor supply



Increasing tax to t_2 raises d, but loses b, while H loses $d + e$

Behavioral responses with upward sloping labor supply



$$MU_L^* \cdot (d-b) > MU_H^* \cdot (d+e)? \Leftrightarrow \frac{MU_L}{MU_H} > \frac{d+e}{d-b} = \frac{d-b+b+e}{d-b} = 1 + \frac{b+e}{d-b}$$

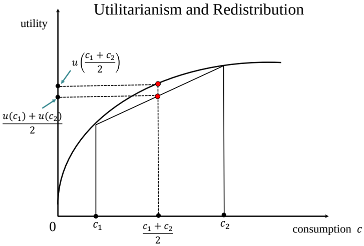
Interpreting the example

$\frac{b+e}{d-b}$ is the ratio of marginal EB to marginal revenue:

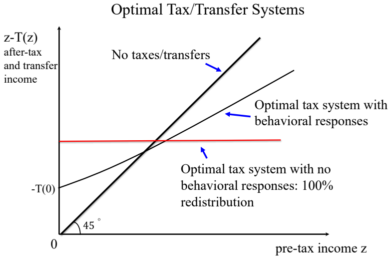
$$\frac{MU_L}{MU_H} > \underbrace{1 + \frac{MEB}{MR}}_{\text{Marginal cost of public funds}}$$

- MCPF captures the tradeoff between equity and efficiency
- If \$1 to L makes H \$1.50 worse off, L's MU must be 1.5x H's lost MU to justify the tax
- The higher ratio of MEB to MR, the higher MU_L to justify the tax
- With diminishing MU and increasing MEB, this is eventually false as incomes converge

No behavioral responses (illustration)



(a) Utilitarian SWF



(b) Optimal tax schedule

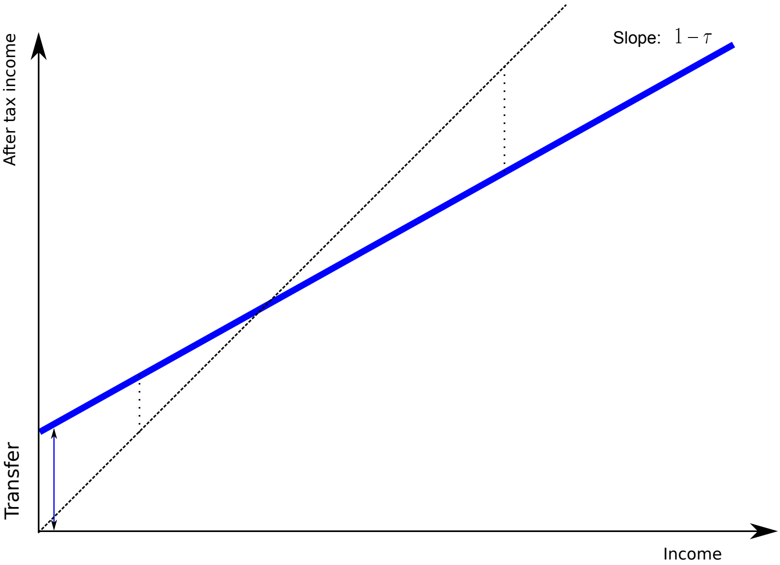
Without behavioral responses, $MEB = 0$ and $\frac{MU_L}{MU_U} > 1$ for all $z_L < z_H$, so a 100% tax is optimal.

More complex optimal *linear* income taxation assumptions

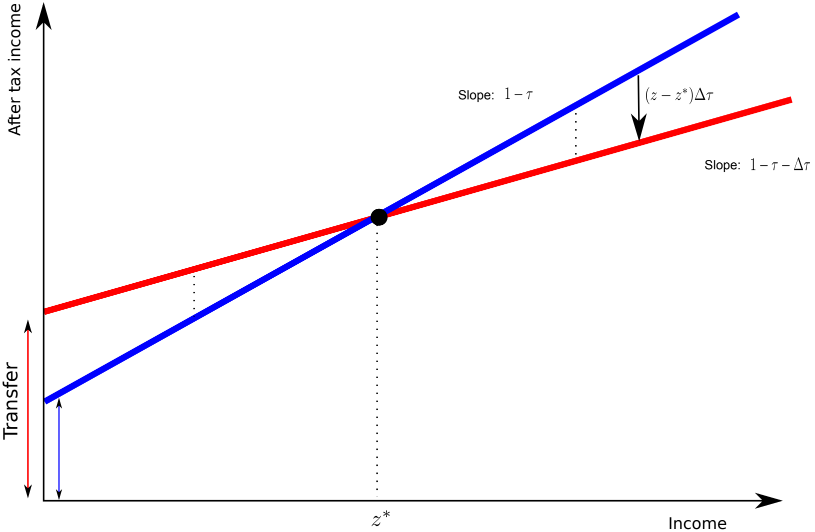
- Linear tax t on income z with uniform lump sum transfer G

[2] Mapping an WF $\omega(u(C, z; w))$ to $\lambda_i \times u(C, z; w)$ is a non-trivial step.

Marginal tax rates and redistribution



Marginal tax rates and redistribution



More complex optimal *linear* income taxation assumptions

- Linear tax t on income z with uniform lump sum transfer G
- The individual's after tax income

$$z - \underbrace{(t \cdot z - G)}_{\text{tax liability}} = G + (1 - t)z$$

[2] Mapping an WF $\omega(u(C, z; w))$ to $\lambda_i \times u(C, z; w)$ is a non-trivial step.

More complex optimal *linear* income taxation assumptions

- Linear tax t on income z with uniform lump sum transfer G
- The individual's after tax income

$$z - \underbrace{(t \cdot z - G)}_{\text{tax liability}} = G + (1 - t)z$$

- Assume utility given consumption C and labor supply L:

$$C - v(L)$$

[2] Mapping an WF $\omega(u(C, z; w))$ to $\lambda_i \times u(C, z; w)$ is a non-trivial step.

More complex optimal *linear* income taxation assumptions

- Linear tax t on income z with uniform lump sum transfer G
- The individual's after tax income

$$z - \underbrace{(t \cdot z - G)}_{\text{tax liability}} = G + (1-t)z$$

- Assume utility given consumption C and labor supply L:

$$C - v(L)$$

- N individuals ($i = 1, \ldots, N$); paid wage (=ability), w_i; $z = w_i L$

$$u(C, z; w_i) = C - v\left(z/w_i\right)$$

[2] Mapping an WF $\omega(u(C,z;w))$ to $\lambda_i \times u(C,z;w)$ is a non-trivial step.

More complex optimal *linear* income taxation assumptions

- Linear tax t on income z with uniform lump sum transfer G
- The individual's after tax income

$$z - \underbrace{(t \cdot z - G)}_{\text{tax liability}} = G + (1-t)z$$

- Assume utility given consumption C and labor supply L:

$$C - v(L)$$

- N individuals ($i = 1, \ldots, N$); paid wage (=ability), w_i; $z = w_i L$

$$u(C, z; w_i) = C - v\left(z/w_i\right)$$

- Gov't maximizes welfare using weights λ_i derived from a concave SWF[2] without running a deficit

$$N \cdot G = t \sum_{i=1}^{N} z_i$$

[2] Mapping an WF $\omega(u(C,z;w))$ to $\lambda_i \times u(C,z;w)$ is a non-trivial step.

Full problem

- We want to pick t and G to maximize

$$\max_{t,G} \sum_{i=1}^{N} \lambda_i u(C - v(z/w_i))$$

Full problem

- We want to pick t and G to maximize

$$\max_{t,G} \sum_{i=1}^{N} \lambda_i u(C - v(z/w_i))$$

- Subject to

Full problem

- We want to pick t and G to maximize

$$\max_{t,G} \sum_{i=1}^{N} \lambda_i u(C - v(z/w_i))$$

- Subject to
  - Individual optimization: given t and G, person i solves

$$\max_{C,z} C - v(z/w_i)$$

  subject to

$$C = G + (1-t)z$$

Full problem

- We want to pick t and G to maximize

$$\max_{t,G} \sum_{i=1}^{N} \lambda_i u(C - v(z/w_i))$$

- Subject to
  - Individual optimization: given t and G, person i solves

$$\max_{C,z} C - v(z/w_i)$$

    subject to

$$C = G + (1-t)z$$

  - Government budget constraint

$$N \cdot G = \sum_{i=1}^{N} t \cdot z_i$$

    where z_i is income of person i

Optimal linear income tax formula

The optimal tax is characterized by

$$\frac{t}{1-t} = -\frac{\text{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}}$$

where

- λ_i is marginal SWF of person i[3]
- ε_i is the elasticity of income with respect to taxation
- z^M is the mean income
- Covariance[4] is negative: marginal welfare falls as income increases, so RHS is positive
- RHS $\uparrow \Rightarrow t \uparrow \Rightarrow$ more progressivity

[3] Involves a somewhat complex normalization – see appendix for full derivation.

[4] Covariance is a measure of how much two variables change together. If negative, they move in opposite directions.

Optimal linear income tax formula

$$\frac{t}{1-t} = -\frac{\mathrm{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}}$$

Interpretation

- higher t is associated with higher progressivity

Optimal linear income tax formula

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}}$$

Interpretation

- higher t is associated with higher progressivity
- optimal progressivity $\uparrow$ if marginal social welfare $\downarrow$ with income ($\uparrow$ covariance), (less welfare loss from taxing them)
- optimal progressivity $\downarrow$ if behavioral response large (ε higher)

Optimal linear income tax formula

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}}$$

Interpretation

- higher t is associated with higher progressivity
- optimal progressivity $\uparrow$ if marginal social welfare $\downarrow$ with income ($\uparrow$ covariance), (less welfare loss from taxing them)
- optimal progressivity $\downarrow$ if behavioral response large (ε higher)
- high-income response to taxes matters more than low-income

Optimal linear income tax formula

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}}$$

Interpretation

- higher t is associated with higher progressivity
- optimal progressivity $\uparrow$ if marginal social welfare $\downarrow$ with income ($\uparrow$ covariance), (less welfare loss from taxing them)
- optimal progressivity $\downarrow$ if behavioral response large (ε higher)
- high-income response to taxes matters more than low-income
- $\uparrow$ redistributive tastes lead to $\uparrow$ progressivity (λ varies more)

Optimal linear income tax formula

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}}$$

Interpretation

- higher t is associated with higher progressivity
- optimal progressivity $\uparrow$ if marginal social welfare $\downarrow$ with income ($\uparrow$ covariance), (less welfare loss from taxing them)
- optimal progressivity $\downarrow$ if behavioral response large (ε higher)
- high-income response to taxes matters more than low-income
- $\uparrow$ redistributive tastes lead to $\uparrow$ progressivity (λ varies more)
- $\uparrow$ inequality means $\uparrow$ progressivity (z_i/z^M varies more)

Connection to Laffer curve

Sometimes rewritten as:

$$\frac{t}{1-t} = -\frac{\text{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}} \iff t = -\frac{1-\bar{g}}{1-\bar{g}+\bar{\varepsilon}_i}$$

where:

- $\bar{g} = \mathbb{E}[\lambda_i \times \frac{z_i}{z^M}]$ - relevant part of covariance formula
- $\bar{\varepsilon} = \mathbb{E}[\varepsilon_i \times \frac{z_i}{z^M}]$ - income-weighted of elasticity

Maps neatly to the Laffer curve:

- If Rawlsian, $\bar{g} = 0$, so $t = \frac{1}{1+\bar{\varepsilon}}$, the Laffer curve!
- If (extreme) "Libertarian," $\bar{g} = 1$[5], so $t = 0$, no tax!
- If Utilitarian, $\bar{g} \in (0,1)$, so $0 < t < \frac{1}{1+\bar{\varepsilon}}$

[5] No social welfare at all, so $-cov(\lambda_i, \frac{z_i}{z^M}) = 1 - 1 = 0$

Sufficient statistics approach

- Some of the optimal tax formula is measurable like income elasticity and income inequality

Sufficient statistics approach

- Some of the optimal tax formula is measurable like income elasticity and income inequality
- Social welfare function is a policy choice

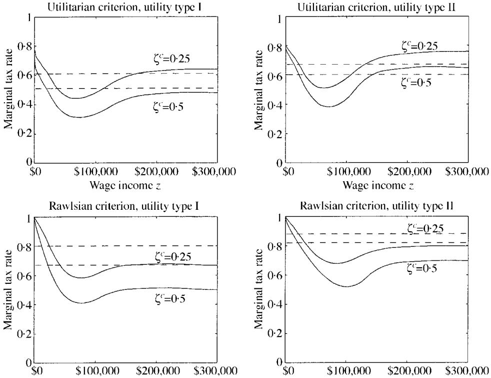
Sufficient statistics approach

- Some of the optimal tax formula is measurable like income elasticity and income inequality
- Social welfare function is a policy choice
- Given SWF, the measurable components are “sufficient” statistics

Sufficient statistics approach

- Some of the optimal tax formula is measurable like income elasticity and income inequality
- Social welfare function is a policy choice
- Given SWF, the measurable components are “sufficient” statistics
- Sufficient statistics separate the positive from the normative

Welfare function matters



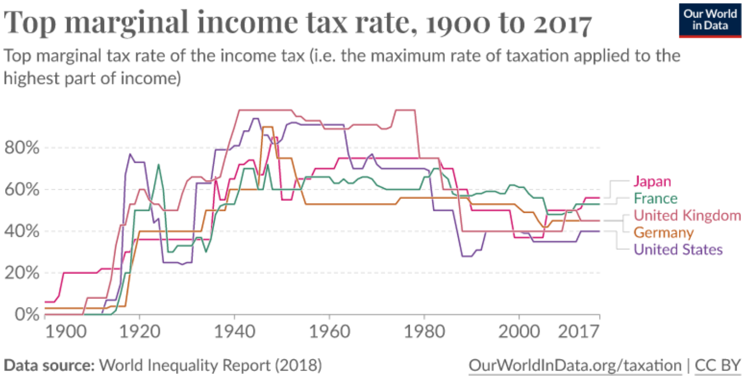
Optimal tax schedules based on income distribution in 1993. Panels vary by utility and social welfare function. Tax rates are high at the bottom due to phase out of transfers. (Saez (2001) Review of Economic Studies)

Nonlinear income tax: top income tax rate

- Optimal top marginal tax rate?

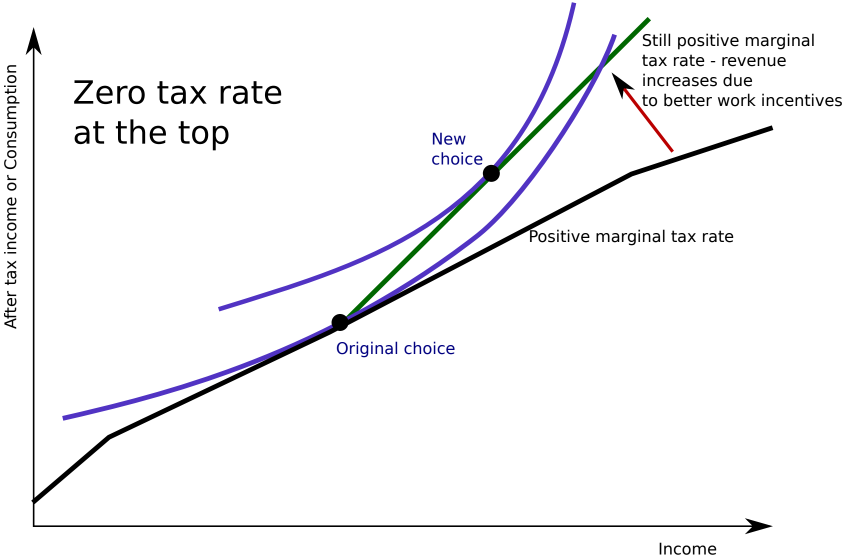
[6] Assumes no utility or income effect for highest earners.

Rates at the top over time



From Our World In Data (Piketty's work)

Nonlinear income tax: zero tax rate at the top



Nonlinear income tax: top income tax rate

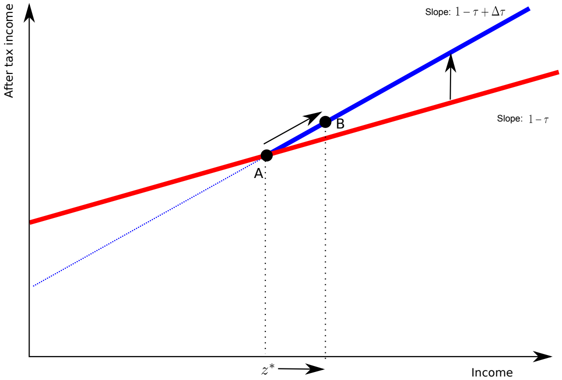
- Optimal top marginal tax rate?
- If one person had the highest ability, optimal rate is zero to maximize work

[6] Assumes no utility or income effect for highest earners.

Wait, what?

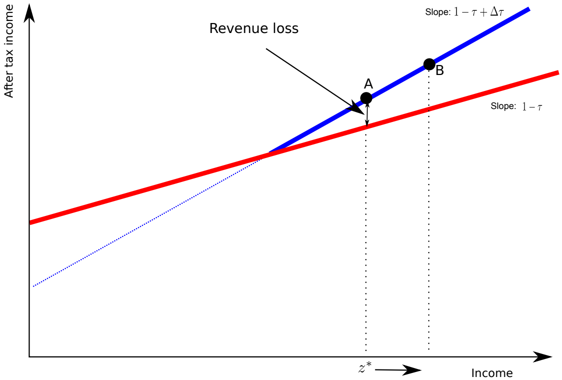
Taken from the Charlotte Observer

Change in top tax rate: Incentive effect



Mechanical increase $A \rightarrow B$ from incentive: $\partial M = [z - z^*]\partial\tau$

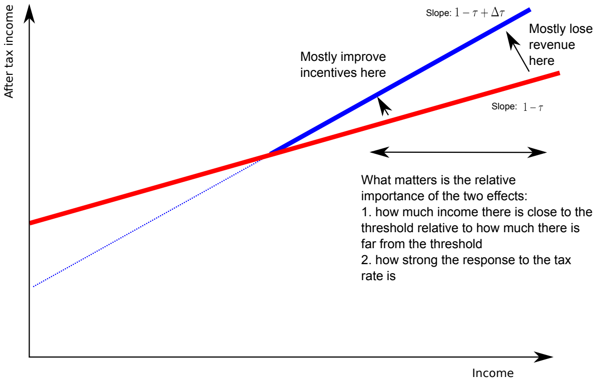
Change in top tax rate: Revenue cost



Revenue cost (gap between lines) from behavioral response:

$$\partial B = \tau \partial z = -\tau \frac{\partial z}{\partial(1-\tau)} \partial \tau = -\frac{\tau}{1-\tau} \varepsilon \cdot z \cdot \partial \tau$$

Non-zero tax rate at the top



Goal: $\partial M + \partial B = \partial\tau \left([z - z^*] - \frac{\tau}{1-\tau}\varepsilon \cdot z \right) = 0$ occurs at

$$\frac{\tau}{1-\tau} = \frac{1}{\varepsilon} \cdot \frac{z - z^*}{z} \Rightarrow \tau = \frac{1}{1 + a\varepsilon} \quad \text{where } a = \frac{z}{z - z^*}$$

Nonlinear income tax: top income tax rate

- Optimal top marginal tax rate?
- If one person had the highest ability, optimal rate is zero to maximize work
- Rarely a single identifiable person, so the top rate depends on the income distribution

[6] Assumes no utility or income effect for highest earners.

Nonlinear income tax: top income tax rate

- Optimal top marginal tax rate?
- If one person had the highest ability, optimal rate is zero to maximize work
- Rarely a single identifiable person, so the top rate depends on the income distribution
- Formula for top tax rate (note similarity to Laffer formula)[6]

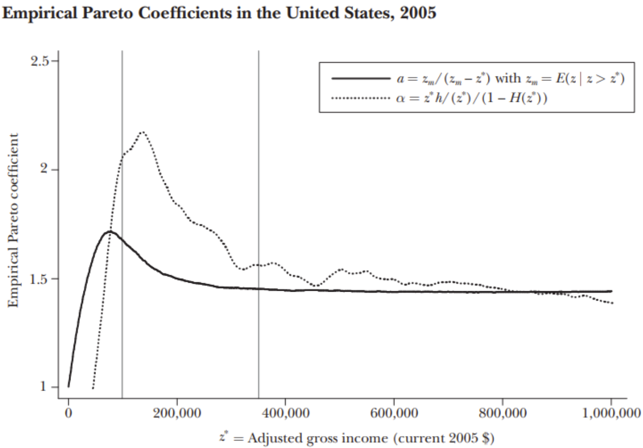
$$\tau = \frac{1}{1 + a\varepsilon}$$

where:

- ε: elasticity of taxable income
- $a = \frac{z}{z - z^*}$: Pareto coefficient measures income over z^*
- Empirically: $a = 1.5$ to 2, $\varepsilon = 0.2$ to $1.0 \Rightarrow \tau \in [71\%, 76\%]$

[6] Assumes no utility or income effect for highest earners.

Pareto coefficient



Average income z^m above a threshold. As z^m increases, a approaches 1. As z^m nears z^*, a becomes very large. A lower a indicates more income inequality and a higher top tax rate. From Saez and Diamond (2011)

Conclusion

- Optimal income taxation is about targeting the ability to pay if observing income
- Degree of progressivity driven in part by observable characteristics like income inequality and in part by subjective characteristics like social welfare function
- Elasticity of income with respect to taxation summarizes behavioral responses to taxation
- Sufficient statistics approach cleanly separates optimal taxation into objective measures and subjective choices

Appendix: Individual optimization

$$\max_{C,z} \; C - v(z/w_i)$$

subject to

$$C = G + (1-t)z$$

- Equivalently:

$$\max_{z} G + (1-t) \cdot z - v(z/w_i)$$

Appendix: Individual optimization

$$\max_{C,z} \; C - v(z/w_i)$$

subject to

$$C = G + (1-t)z$$

- Equivalently:

$$\max_z G + (1-t)\cdot z - v(z/w_i)$$

- Optimum: $v'(z/w_i) = (1-t)w_i$. The solution is $z(t, w_i)$

Appendix: Individual optimization

$$\max_{C,z} \; C - v(z/w_i)$$

subject to

$$C = G + (1-t)z$$

- Equivalently:

$$\max_{z} G + (1-t)\cdot z - v(z/w_i)$$

- Optimum: $v'(z/w_i) = (1-t)w_i$. The solution is $z(t, w_i)$
- Effect of a change in t on utility (envelope theorem):

$$\left((1-t) - v'(z(t,w_i)/w_i)\Big/ w_i\right)\frac{\partial z}{\partial t} - z(t,w_i) = -z(t,w_i)$$

Appendix: Individual optimization

$$\max_{C,z} \; C - v(z/w_i)$$

subject to

$$C = G + (1-t)z$$

- Equivalently:

$$\max_z G + (1-t) \cdot z - v(z/w_i)$$

- Optimum: $v'(z/w_i) = (1-t)w_i$. The solution is $z(t, w_i)$
- Effect of a change in t on utility (envelope theorem):

$$\left((1-t) - v'(z(t, w_i)/w_i) \Big/ w_i\right) \frac{\partial z}{\partial t} - z(t, w_i) = -z(t, w_i)$$

Appendix: Individual optimization

$$\max_{C,z} \; C - v(z/w_i)$$

subject to

$$C = G + (1-t)z$$

- Equivalently:

$$\max_z G + (1-t)\cdot z - v(z/w_i)$$

- Optimum: $v'(z/w_i) = (1-t)w_i$. The solution is $z(t, w_i)$
- Effect of a change in t on utility (envelope theorem):

$$\left((1-t) - v'(z(t,w_i)/w_i)\Big/ w_i\right)\frac{\partial z}{\partial t} - z(t,w_i) = -z(t,w_i)$$

- Effect of a change in G on utility (envelope theorem):

$$\left((1-t) - v'(z(t,w_i)/w_i)\Big/ w_i\right)\frac{\partial z}{\partial G} + 1 = 1$$

Government's problem

- From the revenue constraint, $G(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} t \cdot z(t, w_i)$.

Government's problem

- From the revenue constraint, $G(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} t \cdot z(t, w_i)$.
- $G'(t) = \frac{1}{N} \sum_{i=1}^{N} \left(z(t, w_i) + t \frac{\partial z}{\partial t} \right)$

Government's problem

- From the revenue constraint, $G(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} t \cdot z(t, w_i)$.
- $G'(t) = \frac{1}{N} \sum_{i=1}^{N} \left(z(t, w_i) + t \frac{\partial z}{\partial t} \right)$
- We want to maximize

$$\sum_{i=1}^{N} \omega \Big(G(t) + (1-t) z(t, w_i) - v(z(t, w_i)/w_i) \Big)$$

with respect to t

Government's problem

- From the revenue constraint, $G(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} t \cdot z(t, w_i)$.
- $G'(t) = \frac{1}{N} \sum_{i=1}^{N} \left(z(t, w_i) + t \frac{\partial z}{\partial t} \right)$
- We want to maximize

$$\sum_{i=1}^{N} \omega\Big(G(t) + (1-t) z(t, w_i) - v(z(t, w_i)/w_i) \Big)$$

  with respect to t
- The solution has to satisfy

$$0 = \sum_{i=1}^{N} \omega_i' \cdot \left(G'(t) - z(t, w_i) \right) = N \cdot G'(t) - \sum_{i=1}^{N} \lambda_i I(t, w_i)$$

  where ω_i is the marginal welfare of person i (the value of an extra util; here, equivalent to an extra dollar) and $\lambda_i = \frac{\omega_i'}{\sum_{i=1}^{N} \omega_i' \Big/ N}$ are the normalized welfare weights ($\frac{\sum \lambda_i}{N} = 1$)

Optimum:

- Let's substitute in $G'(t) = \frac{1}{N}\sum_{i=1}^{N}\left(z(t, w_i) + t\frac{\partial z}{\partial t}\right)$

$$N \cdot G'(t) - \sum_{i=1}^{N} \lambda_i I(t, w_i) = \sum_{i=1}^{N}\left(z(t, w_i) + t\frac{\partial z}{\partial t}\right) - \sum_{i=1}^{N} \lambda_i I(t, w_i) = 0$$

Optimum:

- Let's substitute in $G'(t) = \frac{1}{N}\sum_{i=1}^{N}\left(z(t,w_i) + t\frac{\partial z}{\partial t}\right)$

$$N \cdot G'(t) - \sum_{i=1}^{N} \lambda_i I(t,w_i) = \sum_{i=1}^{N}\left(z(t,w_i) + t\frac{\partial z}{\partial t}\right) - \sum_{i=1}^{N} \lambda_i I(t,w_i) = 0$$

- Rearranging (note that $\frac{\partial z}{\partial t} = -\frac{\partial z}{\partial(1-t)}$) and multiplying by $(1-t)/(1-t)$ and z/z:

$$-\frac{t}{1-t} \cdot \sum_{i=1}^{N} \frac{\partial z}{\partial(1-t)} \frac{1-t}{z} \cdot z = \sum_{i=1}^{N} (\lambda_i - 1) z(t,w_i)$$

Optimum:

- Let's substitute in $G'(t) = \frac{1}{N}\sum_{i=1}^{N}\left(z(t,w_i) + t\frac{\partial z}{\partial t}\right)$

$$N \cdot G'(t) - \sum_{i=1}^{N}\lambda_i I(t,w_i) = \sum_{i=1}^{N}\left(z(t,w_i) + t\frac{\partial z}{\partial t}\right) - \sum_{i=1}^{N}\lambda_i I(t,w_i) = 0$$

- Rearranging (note that $\frac{\partial z}{\partial t} = -\frac{\partial z}{\partial(1-t)}$) and multiplying by $(1-t)/(1-t)$ and z/z:

$$-\frac{t}{1-t}\cdot\sum_{i=1}^{N}\frac{\partial z}{\partial(1-t)}\frac{1-t}{z}\cdot z = \sum_{i=1}^{N}(\lambda_i - 1)z(t,w_i)$$

- ...and just one more step:

$$-\frac{t}{1-t}\sum_{i=1}^{N}\varepsilon^{i}\cdot z = \sum_{i=1}^{N}(\lambda_i - 1)\left(z(t,w_i) - z^{M}\right)$$

where average income z^M is added/subtracted and $\frac{\sum \lambda_i}{N} = 1$

Optimal linear income tax formula

The optimal tax is characterized by:

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}} \iff t = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\left(\varepsilon_i \cdot \frac{1}{N}\sum_i \frac{z_i}{z^M}\right) - \operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}$$

where

- $\lambda_i = \frac{\omega_i'}{\sum_{i=1}^{N} \omega_i' \big/ N}$ is marginal SWF of person i normalized to sum to N (makes math easier)
- ε_i is the elasticity of income with respect to taxation
- z^M is the mean income
- Covariance is negative: marginal welfare falls as income increases, so RHS is positive
- As RHS increases, the tax rate t increases meaning more progressivity

Optimal linear income tax formula

The optimal tax is characterized by:

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}} \iff t = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\left(\varepsilon_i \cdot \frac{1}{N}\sum_i \frac{z_i}{z^M}\right) - \operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}$$

where

- $\lambda_i = \dfrac{\omega_i'}{\sum_{i=1}^{N} \omega_i' \Big/ N}$ is marginal SWF of person i normalized to sum to N (makes math easier)
- ε_i is the elasticity of income with respect to taxation
- z^M is the mean income
- Covariance is negative: marginal welfare falls as income increases, so RHS is positive
- As RHS increases, the tax rate t increases meaning more progressivity

Optimal linear income tax formula

The optimal tax is characterized by:

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}} \iff t = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\left(\varepsilon_i \cdot \frac{1}{N}\sum_i \frac{z_i}{z^M}\right) - \operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}$$

where

- $\lambda_i = \frac{\omega_i'}{\sum_{i=1}^{N} \omega_i' \Big/ N}$ is marginal SWF of person i normalized to sum to N (makes math easier)
- ε_i is the elasticity of income with respect to taxation
- z^M is the mean income
- Covariance is negative: marginal welfare falls as income increases, so RHS is positive
- As RHS increases, the tax rate t increases meaning more progressivity

Optimal linear income tax formula

The optimal tax is characterized by:

$$\frac{t}{1-t} = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}} \iff t = -\frac{\operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\left(\varepsilon_i \cdot \frac{1}{N}\sum_i \frac{z_i}{z^M}\right) - \operatorname{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}$$

where

- $\lambda_i = \frac{\omega_i'}{\sum_{i=1}^{N} \omega_i' \big/ N}$ is marginal SWF of person i normalized to sum to N (makes math easier)
- ε_i is the elasticity of income with respect to taxation
- z^M is the mean income
- Covariance is negative: marginal welfare falls as income increases, so RHS is positive
- As RHS increases, the tax rate t increases meaning more progressivity

Optimal linear income tax formula

The optimal tax is characterized by:

$$\frac{t}{1-t} = -\frac{\text{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i \cdot \frac{z_i}{z^M}} \iff t = -\frac{\text{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}{\left(\varepsilon_i \cdot \frac{1}{N}\sum_i \frac{z_i}{z^M}\right) - \text{cov}\left(\lambda_i, \frac{z_i}{z^M}\right)}$$

where

- $\lambda_i = \frac{\omega_i'}{\sum_{i=1}^{N} \omega_i' \big/ N}$ is marginal SWF of person i normalized to sum to N (makes math easier)
- ε_i is the elasticity of income with respect to taxation
- z^M is the mean income
- Covariance is negative: marginal welfare falls as income increases, so RHS is positive
- As RHS increases, the tax rate t increases meaning more progressivity

Aside: Wait how'd you get z^M in there?

Skip this slide if not illuminating, but shows last step for RHS

Aside: Wait how'd you get z^M in there?

Skip this slide if not illuminating, but shows last step for RHS

$$\sum_{i=1}^{N}(\lambda_i - 1)z(t, w_i) = \sum_{i=1}^{N}\lambda_i I(t, w_i) - z(t, w_i)$$

Aside: Wait how'd you get z^M in there?

Skip this slide if not illuminating, but shows last step for RHS

$$\sum_{i=1}^{N}(\lambda_i - 1)z(t, w_i) = \sum_{i=1}^{N}\lambda_i I(t, w_i) - z(t, w_i)$$

$$\textit{Add/subtract } z^M \quad = \sum_{i=1}^{N}\lambda_i I(t, w_i) - z(t, w_i) + z^M - z^M$$

Aside: Wait how'd you get z^M in there?

Skip this slide if not illuminating, but shows last step for RHS

$$\sum_{i=1}^{N}(\lambda_i - 1)z(t, w_i) = \sum_{i=1}^{N} \lambda_i I(t, w_i) - z(t, w_i)$$

$$\textit{Add/subtract } z^M \quad = \sum_{i=1}^{N} \lambda_i I(t, w_i) - z(t, w_i) + z^M - z^M$$

$$\textit{Sum the last } z^M \quad = \sum_{i=1}^{N} (\lambda_i I(t, w_i) - z(t, w_i) + z^M) - N * z^M$$

Aside: Wait how'd you get z^M in there?

Skip this slide if not illuminating, but shows last step for RHS

$$\sum_{i=1}^{N}(\lambda_i - 1)z(t, w_i) = \sum_{i=1}^{N}\lambda_i I(t, w_i) - z(t, w_i)$$

$$\textit{Add/subtract } z^M \quad = \sum_{i=1}^{N}\lambda_i I(t, w_i) - z(t, w_i) + z^M - z^M$$

$$\textit{Sum the last } z^M \quad = \sum_{i=1}^{N}(\lambda_i I(t, w_i) - z(t, w_i) + z^M) - N * z^M$$

$$\textit{Recall } \sum \lambda_i = N \quad = \sum_{i=1}^{N}(\lambda_i I(t, w_i) - z(t, w_i) + z^M) - \sum_{i=1}^{N}\lambda_i * z^M$$

Aside: Wait how'd you get z^M in there?

Skip this slide if not illuminating, but shows last step for RHS

$$\sum_{i=1}^{N}(\lambda_i - 1)z(t, w_i) = \sum_{i=1}^{N}\lambda_i I(t, w_i) - z(t, w_i)$$

$$\textit{Add/subtract } z^M \quad = \sum_{i=1}^{N}\lambda_i I(t, w_i) - z(t, w_i) + z^M - z^M$$

$$\textit{Sum the last } z^M \quad = \sum_{i=1}^{N}(\lambda_i I(t, w_i) - z(t, w_i) + z^M) - N * z^M$$

$$\textit{Recall } \sum \lambda_i = N \quad = \sum_{i=1}^{N}(\lambda_i I(t, w_i) - z(t, w_i) + z^M) - \sum_{i=1}^{N}\lambda_i * z^M$$

$$\textit{“unFOIL”} \quad = \sum_{i=1}^{N}(\lambda_i - 1)(z(t, w_i) - z^M)$$

Aside: Covariance Formula

- There are many summary statistics: mean ($\bar{x}$), median, variance

Aside: Covariance Formula

- There are many summary statistics: mean ($\bar{x}$), median, variance
- Covariance summarizes the joint variability of two variables

Aside: Covariance Formula

- There are many summary statistics: mean ($\bar{x}$), median, variance
- Covariance summarizes the joint variability of two variables
  - If positive, the two variables are increase together

Aside: Covariance Formula

- There are many summary statistics: mean ($\bar{x}$), median, variance
- Covariance summarizes the joint variability of two variables
  - If positive, the two variables are increase together
  - Ig negative, the two variables move in opposite directions

Aside: Covariance Formula

- There are many summary statistics: mean ($\bar{x}$), median, variance
- Covariance summarizes the joint variability of two variables
  - If positive, the two variables are increase together
  - Ig negative, the two variables move in opposite directions
- A statistical formula for covariance:

$$Cov(x, y) = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{N}$$

Aside: Covariance Formula

- There are many summary statistics: mean ($\bar{x}$), median, variance
- Covariance summarizes the joint variability of two variables
  - If positive, the two variables are increase together
  - Ig negative, the two variables move in opposite directions
- A statistical formula for covariance:

$$Cov(x, y) = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{N}$$

- That looks similar to $\sum_{i=1}^{N} (\lambda_i - 1)\Big(z(t, w_i) - z^M\Big)$ noting that mean of λ_i is 1 (by definition) and z^M is mean income