Economics 2450A: Public Economics and Fiscal Policy I

Section 3: Optimal Income Taxation II

Michael Droste

Fall 2022

Outline

1. Review: Optimal Linear Tax Rate
2. Optimal Top Linear Tax Rate
    - Setup
    - Derivation from first-order conditions
    - Derivation with a perturbation approach
3. Optimal Nonlinear Tax Rate: Saez Approach
    - Setup
    - Derivation with a perturbation approach
    - Sub-optimality of negative marginal tax rate
4. Taking Stock

Review: Optimal Linear Income Tax Rate

Optimal Linear Income Tax Rate

- Last week, we reviewed the optimal linear income tax rate.
- Did so in a model with behavioral responses to taxation (taxable income z depends on tax rate τ) and a generalized social welfare function that allowed for arbitrary preferences for redistribution.
- Key result: the optimal linear income tax rate τ satisfies:

$$\tau = \frac{1-\bar{g}}{1-\bar{g}+e} \quad \text{with} \quad \bar{g} \equiv \frac{\int g_i z^i di}{Z \int g_i di} \quad \text{and} \quad g_i \equiv G'(u^i)u_c^i \quad \text{and} \quad e \equiv \frac{\partial Z}{\partial(1-\tau)} \frac{1-\tau}{Z}$$

Interpreting the Optimal Linear Income Tax Rate

- This is a simple-looking formula because it depends on two objects, $\bar{g}$ and e. But interpreting it can be tricky!

- What is $\bar{g}$? Piketty and Saez (2012) give us a few different interpretations, all valid:
    1. (Mike's favorite): the income-weighted average of the social welfare weight g_i.
    2. (Stefanie's favorite): the covariance between g_i and normalized income z^i/Z.
    3. (nobody's favorite): the ratio of the average income weighted by individual social welfare weights g_i to Z.

- What is e? e is the elasticity of taxable income: a 'macro elasticity' that captures how much aggregate taxable income $Z = \int z^i di$ changes as the net-of-tax rate $1 - \tau$ changes. Use net-of-tax rate just so that this quantity is generally non-negative.

Revenue-Maximizing Tax Rate

- Small aside: we could also derive the revenue-maximizing tax rate in our model.
- Differentiating tax revenue R with respect to τ and substituting in the definition of the taxable income elasticity e yielded:

$$\tau^* = \frac{1}{1+e}$$

- Review problem: When does the revenue-maximizing tax rate equal the social welfare-maximizing tax rate? What social welfare function would deliver this equivalence?

Optimal Top Linear Tax Rate

Optimal Top Linear Tax Rate: Motivation

- Suppose we take as given some tax schedule $T(z)$, and our goal is to determine what the optimal linear tax rate should be on incomes exceeding some exogenous amount z^*.

  Example: U.S. tax schedule is piecewise linear, with constant marginal tax rates in several income 'brackets'. What should the top marginal tax rate (currently 37%) be?

- To be very precise and address the Jimmy Critique from lecture, let's suppose there is a mass 1 of (atomistic, i.e. individually measure zero) households with income above z^* and a total mass N of households. Let $q = 1/n$ denote the fraction of households with incoming exceeding z^*.

- We'll suppose that income takes on non-negative values and the income distribution is not bounded: the support of z across all individuals is $[0, \infty)$.

Optimal Top Linear Tax Rate: Notation

- Warning: the notation is a bit loose here – Saez papers are often like this! The ideas and math do go through, we just need to be very careful to understand what each term means. I will try to be more explicit than lecture/Saez and Piketty (2012).

- Let $z(1-\tau)$ denote the average earnings of those above z^* as a function of the net-of-tax rate: $z(1-\tau) = \int_{z^*}^{\infty} z^i(1-\tau, R)h(z)dz$. We're writing it as a function of τ just to remind us it is!

- Review question: We just defined $z(1-\tau)$ to be average earnings for those above z^*, but we could have said total or aggregate earnings for those above z^* as well (they are equivalent). Why? (Hint: previous slide)

Optimal Top Linear Tax Rate: Approach

- We will describe two solution strategies; both yield the same answer.
- Plan A: direct differentiation with first-order conditions
  - Pros: very explicit, just calculus and algebra
  - Cons: a fair bit of mathematical operations, slow to compute
- Plan B: perturbation argument
  - Pros: relatively simple and fast derivation, easy to interpret economic intuition
  - Cons: perhaps not as obvious the first time you see it
- These are both broadly considered local solution methods; they will both generally work well in the models we consider and we present both for pedagogical reasons.
- On an exam, if you are asked to derive a tax formula, the perturbation approach should be your first choice (faster, less room for mistakes).

Optimal Top Linear Tax Rate: Government Problem

- The government chooses the top linear tax rate τ (which applies to income exceeding the exogenous cutoff z^*) to maximize social welfare, which can be written as:

$$\int_{i:z^i<z^*} G\left[u^i\left(z^i - T(z^i) + q\tau[z - z^*] + \tilde{R}, z^i\right)\right]di + \int_{i:z^i\geq z^*} G\left[u^i\left((1-\tau)[z^i - z^*] + [z^* - T(z^*)] + q\tau[z - z^*] + \tilde{R}, z^i\right)\right]di$$

- It is important to be able to interpret each of these terms!

- $\tilde{R}$ denotes all exogenous income that is not due to τ. This could be exogenous non-labor income, or from transfers due to the tax schedule $T(z)$ that applies to income less than z^* (not directly impacted by τ).

- The term $q\tau\left[z - z^*\right]$ indicates a lump-sum transfer financed by the income on high-income earners. The tax revenue raised from the top marginal tax rate τ is rebated lump-sum to everyone equally, so we weight by q.

Optimal Top Linear Tax Rate: FOC

- First order condition for τ, after applying the envelope theorem ($\partial z^i / \partial \tau$ terms drop out due to household FOCs):

$$0 = \int_{i:z^i<z^*} G'(u^i)u_c^i \left[q(z-z^*) - \tau q \frac{dz}{d(1-\tau)} \right] di + \int_{i:z^i \geq z^*} G'(u^i)u_c^i \left[-(z^i - z^*) + q(z-z^*) - \tau q \frac{dz}{d(1-\tau)} \right] di$$

  where the first integral on the RHS corresponds to persons with income below z^*, and the second integral corresponds to those with incomes above z^*.

- Let's take a moment to make sure we understand what these terms communicate; in particular, the terms inside the square brackets within each integral.

Optimal Top Linear Tax Rate: FOC

- Note that the FOC on the previous slide does not match what's in the lecture notes, though they are algebraically equivalent.

- I can rewrite this to match the lecture slides by noting that the $q(z - z^*) - \tau q \frac{dz}{d(1-\tau)}$ terms are in both integrals. Intuitively, a change in the tax rate will change the lump-sum payment to consumers from the change in tax revenue, and this mechanical effect hits everyone, above and below z^*.

$$0 = \int_i G'(u^i)u_c^i \left[q(z - z^*) - \tau q \frac{dz}{d(1-\tau)} \right] di + \int_{i:z^i \geq z^*} -G'(u^i)u_c^i \left[z^i - z^* \right] di$$

Optimal Top Linear Tax Rate: FOC

- As Stefanie said in lecture, we can't just leave terms like $\frac{dz}{d(1-\tau}$ in this FOC - that would be very uncool. And because we are cool, we will instead rewrite the expression in terms of the taxable income elasticity $e = dz/d(1-\tau) \cdot (1-\tau)/z$.

- Let's also substitute in the marginal social welfare weight $g_i = G'(u^i)u^i_c$, since that will simplify things.

- The FOC becomes:

$$0 = \int_i g_i \left[q(z - z^*) - \tau q \frac{dz}{d(1-\tau)} \right] di - \int_{i:z^i \geq z^*} g_i \left[z^i - z^* \right] di$$

Optimal Top Linear Tax Rate

- The rest of the derivation is nearly identical to the algebra that we did for the optimal linear tax rate (previous slide), with one exception - defining $a = z/(z - z^*)$ and substituting this in to get rid of the $z - z^*$ terms.

- Some algebra yields:

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

  where a is the Pareto parameter defined as above, $\bar{g} = \frac{\int g_i [z^i - z^*] didi}{q[z - z^*] \int g_i di}$ is the average income-weighted marginal social welfare weight *for top earners*, and e is the elasticity of taxable income.

- Review question: How can we interpret a? What does it say about the shape of the income distribution (above z^*) if a is high or low?

Optimal Top Linear Tax Rate: Perturbation

- This is great, but it's a lot of work! Wouldn't it be nice if there was an easier way to derive the optimal top linear tax rate in this model?

- Fortunately, there is - using a perturbation-based approach. The idea is to suppose that we "perturb" an existing tax schedule by considering a small reform $d\tau$ to the top rate, and ask what happens to three things:
  1. Mechanical impact, dM: what is the mechanical impact of the tax change on tax revenue? (Review problem: What is dM?)
  2. Welfare impact, dW: What is the welfare impact of the mechanical change in tax revenue? (Review problem: What is dW? (Hint: use $\bar{g}$)
  3. Behavioral response, dB: what is the behavioral response to the tax change in terms of tax revenue?
     (Review problem: What is dB?)

Optimal Top Linear Tax Rate: Perturbation Solution

- At an optimum τ, a small tax reform $d\tau$ is such that $dM + dW + dB = 0$.
- If not, τ cannot have been optimal. For example, suppose $dM + dW + dB > 0$; the tax revenue generated by a tax increase $d\tau > 0$ would more than offset the welfare loss, this reform is welfare-improving and τ cannot be optimal.
- Setting $dM + dW + dB = 0$ and solving for τ is just algebra:

$$\begin{aligned}
0 &= dM + dW + dB \\
&= [z - z^*]d\tau - \bar{g}[z - z^*]d\tau - \frac{\tau}{1-\tau} ezd\tau \\
&= d\tau \left[(1-\bar{g})[z - z^*] - \frac{\tau}{1-\tau} ez \right] \\
\implies \tau &= \frac{1-\bar{g}}{1-\bar{g}+a \cdot e}
\end{aligned}$$

where $a = z/(z - z^*)$, $\bar{g}$ and e defined as before.

Optimal Top Linear Tax Rate

- As with last week's optimal linear tax model, it is important that you understand how to solve this both ways.
- On an exam, if you are asked to derive an optimal tax formula, the perturbation approach will likely be the easiest approach - it's faster and involves less computation.
- There is an alternative approach to the perturbation considered here (for the optimal top linear tax rate) in Piketty and Saez (2012) - largely similar, but does not explicitly invoke dM, dW, dB and relies on 'aggregating up' individual responses. You're welcome to use that style of argument if you find it more comfortable.
- We will continue to see more applications of a perturbation to solve for optimal policies in this class, and it should encourage you that almost all of the applications follow essentially the same steps and argument.

Optimal Nonlinear Tax (Saez)

Optimal Nonlinear Tax (Saez))

- Now let's think about the optimal nonlinear tax using the Saez (2001) approach.

- We will revisit this in the next week through the lens of the Mirrlees model, which is more structural.

- The key virtues of the Saez approach are that:
  1. Solving it is relatively easy: Can use exactly the same perturbation arguments we made in the previous model
  2. The formula is expressed in terms of sufficient statistics.

- We will gain an appreciation for these virtues as we turn to the Mirrlees model in the next couple classes.

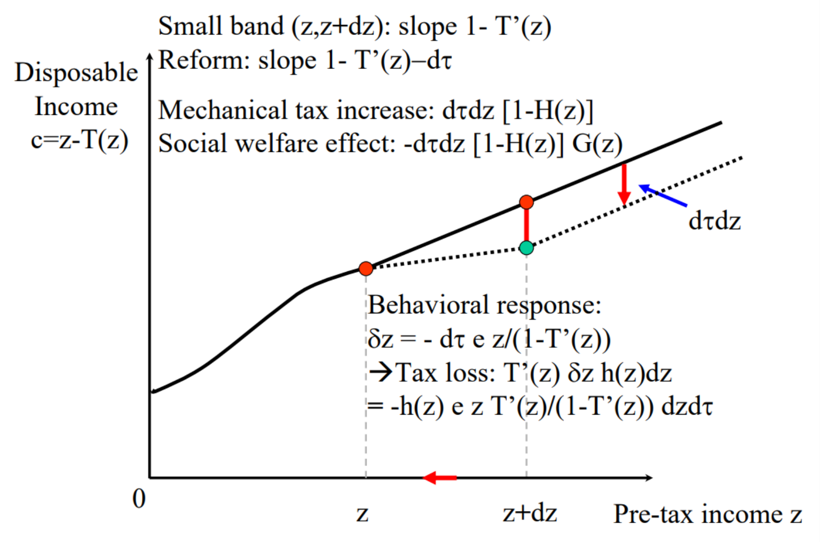
Optimal Nonlinear Tax (Saez): Approach

- The Saez approach to the optimal nonlinear tax is clever, and has the nice graph on the following page.

- The idea: take as given take schedule $T(z)$. For each point z, consider perturbing the tax schedule in an income band $(z, z + dz)$, with dz some very small number, by reducing the marginal tax rate $T'(z)$ by $d\tau$. We assume marginal tax rates are unchanged outside $(z, z + dz)$.

- Denote the income CDF as $H(z)$, the pdf as $h(z)$ (endogenous to tax policy), and let $g(z)$ denote the marginal social welfare weight, $g(z) = G'(u) \cdot u_c / \lambda$. Assume that there are no income effects (quasilinear utility for households); this ends up implying $\int g(z)h(z)dz = 1$.

Optimal Nonlinear Tax (Saez): Key Figure



Optimal Nonlinear Tax (Saez): Notation

- We need to slightly change our notation to accommodate this model, although all of the concepts will remain the same.

- Denote the cumulative distribution of taxable income as $H(z)$ and its associated density function $h(z)$, noting that these are generally endogenous to taxes (income responds to taxation).

- Let $g(z)$ denote the marginal social welfare weight, $g(z) = G'(u) \cdot u_c/\lambda$. Assume that there are no income effects (quasilinear utility for households); this ends up implying $\int g(z)h(z)dz = 1$.

- Let $G(z)$ denote the average social marginal value of consumption for taxpayers with income above z: $G(z) = \frac{\int_z^\infty g(s)h(s)ds}{1-H(z)}$

Optimal Nonlinear Tax (Saez): Perturbation

- I find it useful to stare at the graph while considering dM, dW, dB for a model like this. You may want to draw it out!

- Mechanical impact: Raising the marginal tax rate by $d\tau$ in the band $[z, z + dz]$ increases the tax burden for *everyone* with incomes above z; that is, a mass $1 - H(z)$. Hence, $dM = dz \cdot d\tau \cdot [1 - H(z)]$

- Welfare impact: This is always the easy part! In welfare terms, $dW = -G(z) \cdot dM = -dz \cdot d\tau \cdot [1 - H(z)]G(z)$.

- Behavioral response: A bit more complicated, easy to split in two (next slide).

Optimal Nonlinear Tax (Saez): Perturbation

- Behavioral response: Individuals in the income band $[z, z + dz]$ have a behavioral response, which for convenience we can denote $\delta(z)$ for now (more in a bit).

- Taking as given $\delta(z)$, we can write the behavioral response as $dB = h(z) \cdot dz \cdot T'(z) \cdot \delta(z)$ – the behavioral response is equal to the mass of people at the band, $h(z) \cdot dz$ times the marginal tax rate times the behavioral response.

- So what's $\delta(z)$? Pretty much the same term as in the top linear tax rate: $\delta(z) = -d\tau \cdot e \cdot z/(1 - T'(z))$.

Optimal Nonlinear Tax (Saez): Optimal T'

- Next, define a *local* Pareto parameter at z to be $a(z) = zh(z)/1 - H(z)$. Plays the same role as a did in top linear tax rate model!
- Also define $e(z) = \frac{\partial z}{\partial(1-T'(z))} \cdot \frac{1-T'(z)}{z}$ as the local taxable income elasticity at z. Likewise, this is analogous to our e from before, but now the marginal tax rate varies with z.
- Setting $dM + dW + dB = 0$ and solving for T' yields the optimal (nonlinear) marginal tax rate for any given z:

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

- Very nice: everything generalizes! The optimal marginal tax rate at any income level z depends on the Pareto parameter $a(z)$, the taxable income elasticity $e(z)$, and the average social marginal value of consumption for those with incomes above z, $G(z)$.

Aside: Negative Marginal Tax Rates

- We can use our setup to demonstrate that negative marginal tax rates are generally suboptimal in the simple nonlinear tax model.

- Suppose we have $T'(z) < 0$ for some z, and consider a reform in a band $[z, z + dz]$ (for small dz) that increases the marginal tax rate (brings closer to 0) by a small amount $d\tau$

- Perturbation logic: figure out dM, dW, dB.

  dM: mechanical impact positive (revenue goes up)

  dW: negative, but $dW = -dM \cdot G(z)$, so $dM + dW = dM(1 - G(z)) > 0$ since $1 - G(z) > 0$.

  dB: positive!! This relies on T' being negative to start with (why?)

  $\implies$ Reform $d\tau$ always has $dM + dW + dB > 0$, and so $T'(z) < 0$ not optimal.

Taking Stock

- We've now solved a bunch of optimal income tax models.
- Lucky for us, the optimal tax policies all share similar formulas and reflect similar equity-efficiency trade-offs.
- The approaches we've considered so far exploit the envelope theorem and consider the impacts of small reforms, where the mechanical impacts of the tax are all that appears.
- The key objects:
  - Elasticity of taxable income (sufficient for a revenue-maximizing tax policy)
  - Redistributive preferences, captured by the social welfare function and the (endogenous) distribution of income
  - The shape of the income distribution (e.g. the Pareto tail parameter, if income is Pareto)

Next Week

- Two important objectives for next week:
  1. Important extensions of Saez framework (e.g. tax avoidance, international migration)
  2. Mirrlees model of optimal nonlinear taxation
- See you then!