Public Economics: Lecture 14
Optimal Income Taxation

Cameron LaPoint[‡]

August 2, 2017

In previous lectures, we focused on the positive implications (i.e. the consumer response) of introducing various tax and transfer systems. Now we instead turn to answering normative questions – what kind of taxes *should* the government implement? Here we study what it means for a tax to be optimal, derive the optimal tax formula for the linear income tax problem, and discuss the economic significance of the marginal cost of funds.

1 Optimal Taxation Problems: General Principles

What kind of taxes the government *should* implement depends on two main considerations: (a) the goals of the government, and (b) what tax instruments are available. In general, the government wishes to achieve redistribution of income from one group of citizens to another group via taxes and transfers. The optimal commodity taxation problem we analyzed in previous lectures featured a government that wanted to generate the highest level of utility for its citizens as possible, but could only do so by taxing the consumption of a couple goods available for purchase in the economy. Specifically, in that problem lump-sum tax/transfers were not available, and one good could not be taxed at all. However, in the optimal commodity tax problem, since all taxpayers were assumed to be identical there was no redistribution motive.

Once we know the government's preferences for redistribution and the tax instruments it can implement, two additional considerations come into play: (c) the government must raise a certain amount of revenue, and (d) imposing taxes may affect behavior of those subject to the tax. Point (c) defines the government's budget constraint. In the problems we will study in this course the government always runs a balanced budget – it spends on its citizens exactly as much as it collects in taxes (although in practice this is rarely how governments operate).

Point (d) is the main conceptual issue. People in the economy take taxes and transfers as given and maximize their utility by making decisions that are within their control (i.e.

[‡]Please email me at cl3280@columbia.edu should you find any typos or errors. Any comments or suggestions for improvement are appreciated.

they pick consumption and labor supply). The government must choose taxes that not only balance the budget, but that also take into account how each person will respond to the tax it sets. The government faces a tradeoff between realizing its goals (a) and distortions from altering the optimal decision of its citizens (d). We will see that this tension manifests itself as an **equity-efficiency** tradeoff.

These principles are best illustrated by working through an example of a simple optimal income taxation problem.

Example 1
Suppose there is only one representative person in the economy with utility function $u(C, L) = C - 2L^2$. Here L denotes labor supply, not leisure. Since working is undesirable, it enters negatively (i.e. as a bad rather than a good) into the utility function. The government levies a proportional earnings tax t, and there are no savings, so the individual simply consumes their earnings from labor.

(i) Solve for optimal consumption and labor supply as a function of the wage rate and t.

The maximization problem here can be written as

$$\max_{C,L} \left\{C - 2L^2\right\} \quad \text{s.t. } C = (1-t)wL$$

We can turn this into an unconstrained problem by substituting the budget constraint into the utility function for C.

$$\max_{L} \left\{(1-t)wL - 2L^2\right\}$$

Taking the derivative with respect to L and setting it equal to zero we get a condition for the solution.

$$(1-t)w - 4L = 0 \implies L^* = (1-t)(w/4)$$

We call this type of condition the **first-order condition** (FOC) of the maximization problem. After substituting our solution for labor supply back into the budget constraint, we get optimal consumption $C^* = (1-t)^2(w^2/4)$.

(ii) Would an increase in the tax rate generate an income effect here?

There would be no income effect because the utility function is *quasi-linear* in C; it is linear in consumption but not in labor supply. This means that the marginal rate of substitution between leisure and consumption is independent of consumption. Hence a tax increase will only generate a labor supply response through the substitution effect (leisure becomes relatively cheaper).

(iii) Compute the maximized level of utility. How does this level of utility change with the tax rate?

To find the maximized utility level, substitute the solution C^* and L^* back into the utility function:

$$U(C^*, L^*) = \frac{1}{4}(1-t)^2w^2 - \frac{1}{8}(1-t)^2w^2 = \frac{1}{8}(1-t)^2w^2$$

Taking the derivative of this maximized utility level with respect to the tax rate we get

$$\frac{\partial U(C^*, L^*)}{\partial t} = -\frac{1}{4}(1-t)w^2 < 0$$

An increase in the tax rate thus lowers the level of utility at the optimum.

(iv) What is government revenue as a function of w and t? How does collected revenue change with the tax rate?

Government revenue is the tax bill collected from the individual who optimizes over C^* and L^*:

$$R = twL^* = t(1-t)(w^2/4) = \frac{1}{4}(tw^2 - t^2w^2)$$

Taking the derivative of collected revenue with respect to the tax rate we obtain

$$\frac{\partial R}{\partial t} = \frac{1}{4}(w^2 - 2tw^2) = \frac{1}{4}(1-2t)w^2$$

This condition shows us that when the government sets $t > 1/2$ revenue falls with a tax increase. In this case labor is disincentivized to such an extent that the higher tax rate on each dollar of income is offset by a reduction in labor supply through the substitution effect.

(v) Suppose the government uses all the revenue it collects from the income tax to provide a transfer to the representative individual. Set up the government's optimal tax problem in this scenario and write it as an unconstrained problem.

The government faces two constraints: 1) it must run a balanced budget, so the amount the government rebates to the consumer must equal the amount of tax it collects, and 2) it must take into account that the consumer will pick C and L to maximize utility. Since there is just one person in this economy, the government picks t to generate the highest possible level of utility. The full problem can then be written compactly as

$$\max_{R,t} \left\{C - 2L^2\right\} \text{ s.t. } R = twL$$
$$\text{and s.t. } \max_{C,L} \left\{C - 2L^2\right\} \text{ s.t. } C = R + (1-t)wL$$

We see that the full program of the government <u>nests</u> the individual consumer's problem. To simplify things, first solve the consumer's problem by substituting the budget constraint into the utility function. This gives us a solution of $L^* = (1-t)(w/4)$ and $C^* = R + (1-t)^2(w^2/4)$.

The next step is to compute revenue the government can collect, given the optimal labor supply choice of the individual: $R = twL^* = t(1-t)(w^2/4)$. Now that we have revenue as a function of only the tax rate, we can substitute in for R in the individual's optimal consumption to get

$$C^* = \frac{1}{4}t(1-t)w^2 + \frac{1}{4}(1-t)^2 w^2$$

Finally, we can characterize the maximum level of utility as a function of t to write the unconstrained problem of the government:

$$\max_t \left\{C^* - 2(L^*)^2\right\} \iff \max_t \left\{\frac{1}{4}t(1-t)w^2 + \frac{1}{8}(1-t)^2 w^2\right\}$$

We can now solve for the optimal tax rate t^* by taking the FOC and solving for t^* (this last step is left as an exercise).

2 The Optimal Linear Income Tax Problem

We now apply the techniques used to solve the previous example towards solving the optimal linear income tax problem. Recall the basic features of the problem:

- There are N individuals in the economy, each with their own skill level that the government cannot observe. Each person's skill level corresponds to a wage rate w_i, for individuals $i = 1, 2, \ldots, N$.

- Each person's income is given by $I_i = w_i \cdot L_i$, where L_i is *labor supply* of individual i. While the government can observe each person's income, it does not know to what extent that income is due to inherent ability captured by w_i, or due to effort or preferences for leisure, as captured by L_i.

- The government can impose a tax on income $T(I)$, and this tax is restricted to be linear, so $0 \leq T'(I) \leq 1$. The government also has a preference for redistribution. That is, it picks a tax rate t to maximize a weighted sum of all the utilities of people in the economy; this is the social welfare function $\mathcal{W}\Big(u(C, I; w)\Big)$.

- In choosing a t to maximize welfare, the government takes into account the fact that each person in the economy will choose consumption and labor supply to maximize his/her individual utility.

- The government runs a balanced budget, so whatever it collects in tax revenue gets rebated in equal portions to each consumer as a lump-sum transfer G. The government's budget is then $NG = t\sum_{i=1}^{N} I_i$.

The Full Problem

Given this environment, the full problem the government faces is:

$$\max_{t,G} \left\{ \sum_{i=1}^{N} \mathcal{W}\Big(u(C, I; w)\Big) \right\}$$

$$\text{s.t. } NG = t\sum_{i=1}^{N} I_i$$

$$\text{s.t. individual maximization}$$

Like the more simple problem studied in Example 1, the government's problem here nests each individual's problem of choosing consumption and labor supply to maximize his/her own utility. This means that whatever tax rate and transfer amount the government chooses, it must take into account the fact that individuals report income optimally.

Individual Optimization

To keep things simple, we assume each person has a quasi-linear utility function $u(C, I; w) = C - v(L_i) = C - v(I/w_i)$. We can then write the individual optimization problem as

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

$$\text{s.t. } C = G + (1-t)I$$

$$\iff \max_{I} \Big\{G + (1-t)I - v(I/w_i)\Big\}$$

As usual, the individual picks consumption and income (or labor supply) to maximize utility subject to a budget constraint. We can write this as an unconstrained problem by substituting the budget constraint in for C. The FOC to this problem is $v'(I/w_i) = (1-t)w_i$. For a particular assumption about the functional form of $v(\cdot)$, we get the individual's optimal level of reported income, $I(t, w_i)$.

Example 2

Although this isn't necessary to solve for the optimal income tax formula, it is worth noting that we can easily compute the effect of a change in t or G on maximized utility by applying the Envelope theorem. By the chain rule we have that

$$\frac{du(I(t,w_i))}{dt} = \frac{\partial u(I(\cdot))}{\partial I}\frac{dI(t,w_i)}{dt} + \frac{\partial u(I(\cdot))}{\partial t} = \frac{\partial u(I(\cdot))}{\partial t} = -I(t,w_i)$$

$$\frac{du(I(t,w_i))}{dG} = \frac{\partial u(I(\cdot))}{\partial I}\frac{dI(t,w_i)}{dG} + \frac{\partial u(I(\cdot))}{\partial G} = \frac{\partial u(I(\cdot))}{\partial G} = 1$$

The **Envelope theorem** says that the first term in each of the above lines is zero – the consumer already maximized utility by choosing income $I(t,w_i)$, so $\frac{\partial u(I(\cdot))}{\partial I} = 0$. Put another way, since each person picks their labor supply (and hence income) to maximize utility, the change in the tax rate or transfer amount can only alter utility directly through the budget constraint, not by indirectly affecting the choice of $I(t,w_i)$.

Simplified Government Problem

Now that we have the level of income that solves each individual's problem, $I(t,w_i)$, we can simplify the government's problem. First we can eliminate the second constraint of the goverment's problem (individual maximization) by plugging in the solution $I(t,w_i)$ into the government revenue constraint. Rearranging the revenue constraint we obtain

$$NG = t\sum_{i=1}^{N} I(t,w_i) \implies G(t) = \frac{1}{N}\sum_{i=1}^{N} t\cdot I(t,w_i)$$

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

where the expression for $G'(t)$ follows from the chain rule. Substituting $G(t)$ and the optimized utility of the individual, $u(I(t,w_i)) = G(t) + (1-t)I(t,w_i) - v(I(t,w_i/w_i)$ into the government's welfare function, we have turned the government's complicated problem with two constraints into an unconstrained problem:

$$\max_t \left\{\sum_{i=1}^{N}\mathcal{W}\Big(G(t) + (1-t)I(t,w_i) - v(I(t,w_i)/w_i)\Big)\right\} \tag{2}$$

The solution has to satisfy the FOC with respect to the tax rate:[1]

$$\sum_{i=1}^{N}\omega_i(t)\cdot\Big(G'(t) - I(t,w_i)\Big) = 0 \tag{3}$$

[1]Note that setting the derivative of (2) with respect to t defines a solution here because the social welfare function $\mathcal{W}(\cdot)$ is concave. The fact that the government likes to redistribute income means that it does not like variation in utilities across individuals. This is similar to the idea that a risk averse person does not like variation in income across different states of the world.

Again we used the chain rule to evaluate this derivative, and the shorthand notation $\omega_i(t) = \partial \mathcal{W}(\cdot)/\partial t$ to emphasize that the derivative of the social welfare function depends on each individual i's income. We call the ω_i terms the marginal welfare of person i, or the value to the government of an extra dollar of income given to person i.

Next we define the concept of a **normalized welfare weight**. This is the marginal welfare of each individual relative to the average marginal welfare among people in the economy. Mathematically each of these weights is equal to:

$$\lambda_i = \frac{\omega_i}{\frac{1}{N}\sum_{i=1}^{N}\omega_i}$$

The important thing about these weights is that they summarize how much the government cares about the individual utility of each person in the economy. If the government has a strong preference for redistribution, these λ_i will be higher for people who are relatively poor.

Using these welfare weights, we can rewrite the FOC of equation (3) as

$$\begin{aligned}
&\sum_{i=1}^{N}\omega_i \cdot G'(t) - \sum_{i=1}^{N}\omega_i \cdot I(t, w_i) = 0 \\
\iff &\sum_{i=1}^{N}\left[\lambda_i \cdot \left(\cancel{\frac{1}{N}\sum_{i=1}^{N}\omega_i}\right)\right]G'(t) - \sum_{i=1}^{N}\left[\lambda_i \cdot I(t, w_i)\left(\cancel{\frac{1}{N}\sum_{i=1}^{N}\omega_i}\right)\right] = 0 \\
\iff &G'(t) \cdot \sum_{i=1}^{N}\lambda_i - \sum_{i=1}^{N}\lambda_i \cdot I(t, w_i) = 0 \\
\iff &NG'(t) - \sum_{i=1}^{N}\lambda_i \cdot I(t, w_i) = 0 \qquad (4)
\end{aligned}$$

where we used the fact that $\sum_i \lambda_i/N = 1$ to get from the third to the fourth line. Now plug expression (1) for $G'(t)$ into (4) to get

$$\sum_{i=1}^{N}\left(I(t, w_i) + t \cdot \frac{\partial I(\cdot)}{\partial t}\right) - \sum_{i=1}^{N}\lambda_i \cdot I(t, w_i) = 0 \qquad (5)$$

Finally, we wish to write this condition for the optimal tax rate in terms of an income elasticity that measures the response of each individual i to changes in the tax rate. To keep the notation manageable, define a few objects:

$$\begin{aligned}
\epsilon^i &= \frac{\partial I_i}{\partial(1-t)}\left(\frac{1-t}{I_i}\right) = -\frac{\partial I_i}{\partial t}\left(\frac{1-t}{I_i}\right) \\
I^M &= \frac{1}{N}\sum_{i=1}^{N} I_i
\end{aligned}$$

The parameter ϵ^i represents the elasticity of income to the retention rate $(1-t)$, and I^M is the average level of reported income in the economy. After a few more rearrangements of expression (5) we end up with:

$$
\begin{aligned}
& -\frac{t}{1-t} \cdot \sum_{i=1}^{N} \frac{\partial I_i}{\partial(1-t)} \frac{1-t}{I_i} \cdot I_i = \sum_{i=1}^{N} (\lambda_i - 1) \cdot I(t, w_i) \\
\iff & -\frac{t}{1-t} \cdot \sum_{i=1}^{N} \epsilon^i \cdot I_i = \sum_{i=1}^{N} (\lambda_i - 1)\Big(I(t, w_i) - I^M\Big)
\end{aligned} \tag{6}
$$

where the last line follows from subtracting a zero because $\sum_{i=1} \lambda_i / N = 1$.[2] The very last step is to rearrange (6) so that the tax progressivity term $t/(1-t)$ is on the LHS, and the RHS features the covariance of the welfare weights with income.[3]

$$
\begin{aligned}
& \frac{t}{1-t} = -\frac{\frac{1}{N}\sum_{i}^{N} (\lambda_i - 1)\Big(I(t, w_i) - I^M\Big)}{\frac{1}{N}\sum_{i=1}^{N} \epsilon^i \cdot I} \\
\iff & \frac{t}{1-t} = -\frac{\text{cov}(\lambda_i, \frac{I_i}{I^M})}{\frac{1}{N}\sum_i \epsilon^i \cdot \frac{I_i}{I^M}}
\end{aligned} \tag{7}
$$

Optimal Progressivity Formula

Equation (7) gives us what is sometimes called the optimal tax progressivity formula. The LHS of this formula gives the marginal tax rate t relative to the marginal retention rate $(1-t)$, whereas the RHS represents an equity-efficiency tradeoff. The numerator term on the RHS represents the government's preferences for redistribution, while the denominator term represents the efficiency cost associated with people's behavioral response to tax changes. A higher t means a more progressive tax policy.

From the RHS of (7), there are three main features that affect the progressivity of the tax:

1. The optimal tax rate is lower if people are highly responsive to taxation. That is, if the elasticity of income with respect to the retention rate (ϵ^i) is high, the term representing inefficiency in the denominator is going to be high. Observe also that the response of people with higher levels of taxable income matters more than those with lower income, because the ϵ^i in the denominator are weighted by each individual i's income relative to the average level of income I_i/I^M.

[2] In particular, $\sum_{i=1}^{N}(\lambda_i - 1)I^M = \sum_{i=1}^{N}(\lambda_i - 1)\Big(\sum_{i=1}^{N} I_i/N\Big) = \sum_{i=1}^{N} I_i - N\sum_{i=1}^{N} I_i/N = 0$.

[3] Here we used the sample covariance formula for any two random variables x, y: $cov(x, y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$, where here $\bar{x} = \frac{1}{N}\sum_i \lambda_i = 1$ and $\bar{y} = \sum_i I_i/I^M = 1$. Don't worry if you don't understand this step.

2. Progressivity of the tax should rise with stronger government preferences for redistribution (the numerator term is high). When tastes for redistribution are higher, the covariance term is more negative. A negative covariance term here means that low values of the welfare weights λ_i correspond to high income levels, indicating that the government cares less about the utility of high-income people.

3. The optimal tax rate should be higher if income is a better indicator of the ability to pay. Recall that in the optimal income tax problem, the government cannot directly observe each person's ability to pay.[4] It instead relies on income as a proxy for ability to pay. If income fluctuates wildly for each person year-to-year (for instance, due to luck), this measure will be a bad indicator of ability to pay. In the formula, when income is a bad indicator the covariance term will be close to zero.

A few final notes on this problem. Unlike the optimal commodity tax problem, we do not assume directly that the government cannot impose a lump sum tax ($T'(I) = 0$ is possible). However, the setup of the problem does restrict marginal tax rates to be positive, $T'(I) \geq 0$, so the optimal tax formula will never feature subsidies.[5]

3 Marginal Cost of Funds (MCF)

The optimal progressivity formula in the income tax problem describes a particular type of equity-efficiency tradeoff. One way to quantify this severity of this tradeoff in units of utility is the **marginal cost of funds** (MCF). Here are three equivalent definitions of the MCF:

- $\text{MCF} = \frac{\text{effect of the tax instrument on welfare}}{\text{effect of the tax instrument on revenue}}$
- $\text{MCF} = -\frac{\partial U^*/\partial t}{\partial R^*/\partial t}$, where U^* is the level of utility that the individual achieves by optimizing, and R^* is revenue the government collects given the individual's optimal choices.
- The MCF is the welfare loss associated with raising an additional dollar of revenue.

Example 3
Let's return to the setup of Example 1 and compute the MCF. Recall in that example we computed the effect of the proportional income tax on welfare (i.e. maximized utility), and the effect of the tax on revenue:

$$\begin{cases} \frac{\partial U^*}{\partial t} = -\frac{1}{4}(1-t)w^2 \\ \frac{\partial R^*}{\partial t} = \frac{1}{4}(1-2t)w^2 \end{cases} \implies \text{MCF} = \frac{(1-t)}{(1-2t)}$$

[4]For this reason the optimal income taxation problem is often thought of as a particular kind of adverse selection problem governments face.

[5]We could allow for portions of the marginal tax rate schedule to be positive or negative by studying a more general problem where taxes as a function of income $T(I)$ can be nonlinear. This is a much more mathematically complex problem to solve, but the solution to the more general problem retains the intuition of an equity-efficiency tradeoff in setting tax rates.

This is the MCF for any tax rate t that the government might set. However, we can show using the Envelope theorem that when the government chooses the optimal tax rate t^*, we have MCF = 1. First note that we can write optimized utility as

$$U(C^*, L^*) = (1-t)wL^* + R^* - 2(L^*)^2$$

Taking the FOC with respect to t and recalling that L^* and R^* depend on t we obtain

$$\frac{\partial U(C^*, L^*)}{\partial t} = \Big[(1-t)w - 4L^*\Big]\frac{\partial L^*}{\partial t} - wL^* + \frac{\partial R}{\partial t} = 0$$

The first term above is equal to zero, since it is equal to $\frac{\partial U^*}{\partial L^*}\frac{\partial L^*}{\partial t}$, and $\frac{\partial U^*}{\partial L^*} = 0$ by the Envelope theorem. After applying this fact, we learn that

$$\frac{\partial R^*}{\partial t} = wL^* = \frac{1}{4}(1-t^*)w^2 = -\frac{\partial U^*}{\partial t}$$

The fact that the MCF = 1 for the optimal tax rate implies that the government sets $t^* = 0$. Starting at this zero tax rate, the welfare loss from a small tax increase is exactly equal to the revenue gain.