Economics 230a, Fall 2021
Lecture Note 3: Further Optimal Tax Results

Distributional Considerations

The basic Ramsey rule is derived under the assumption that we are trying to maximize the utility of a representative individual, so only efficiency considerations matter. Yet to make sense of our inability to use lump-sum taxes, we need some sort of heterogeneity in the population. So, assume that individuals differ in some unspecified manner, and consider an extension of the optimal tax problem where we have the same set of instruments but now seek to maximize social welfare, $W(V^1(\boldsymbol{p})\ V^2(\boldsymbol{p}), \ldots, V^H(\boldsymbol{p}))$, subject to satisfying the revenue constraint that $(\boldsymbol{p} - \boldsymbol{q})'\boldsymbol{X} \geq R$, where $\boldsymbol{X} = \Sigma_h \boldsymbol{x}^h$ is the vector of total consumption by households. Setting up the Lagrangian with μ as the shadow price of the revenue constraint, we obtain the first-order conditions:

$$(1) \qquad -\sum_h W_h \lambda^h x_i^h + \mu\left[X_i + \sum_j t_j \sum_h \frac{dx_j^h}{dp_i}\right] = 0 \qquad \forall i$$

where $W_h\lambda^h$ is the marginal welfare effect of an increase in individual h's income. Once again using the Slutsky equation to break each individual price effect dx_j^h/dp_i into income and substitution effects, and grouping terms, we get:

$$(2) \qquad \left[\mu - \left(\frac{\sum_h x_i^h\left(W_h\lambda^h + \mu\sum_j t_j \frac{dx_j^h}{dy^h}\right)}{X_i}\right)\right] X_i + \mu\sum_j t_j S_{ji} = 0 \ \Rightarrow\ -\sum_j t_j S_{ji} = \frac{\mu - \alpha_i}{\mu} X_i \qquad \forall i$$

where $S_{ji} = \sum_h s_{ji}^h$ is the sum of the Slutsky terms across individuals and we may think of $\alpha_i = \frac{\sum_h x_i^h\left(W_h\lambda^h + \mu\sum_j t_j \frac{dx_j^h}{dy^h}\right)}{X_i} = \frac{\sum_h x_i^h \alpha^h}{X_i}$ as the marginal social welfare of income associated with good i; it equals the average of the social welfare of individual incomes, α^h, weighted by individual shares in good i's consumption, x_i^h/X_i. Recalling that the term $-\sum_j t_j S_{ji}$ equals the marginal excess burden from an increase in the tax on good i, expression (2) implies that the ratio of this excess burden to the revenue associated with good i, $X_i + \sum_j t_j S_{ji}$, should equal $\frac{\mu - \alpha_i}{\alpha_i}$. It is no longer optimal to set the <u>marginal cost of public funds</u> (revenue plus excess burden per unit of revenue) equal for all revenue sources; we now wish to take into account who consumes the goods; for goods with a higher positive correlation between x_i^h/X_i and α^h, α_i will be higher and hence the desired marginal cost of funds should be lower. Relative to the representative agent case, we should lower taxes on goods purchased relatively intensively by those with higher social income weights – presumably those of lower ability and income. As to the overall impact of equity and efficiency considerations, consider again the example with two taxed goods. The modified Ramsey rule in (2) becomes:

$$(3) \qquad \frac{t_1/p_1}{t_2/p_2} = \frac{\pi_1\varepsilon_{20} + \pi_2\varepsilon_{12} + \pi_1\varepsilon_{21}}{\pi_2\varepsilon_{10} + \pi_2\varepsilon_{12} + \pi_1\varepsilon_{21}} \qquad \text{where } \pi_i = \frac{\mu - \alpha_i}{\mu}.$$

As only the first terms in numerator and denominator of (3) differ, the proportional tax on good 1 will now be higher than the tax on good 2 if and only if $\varepsilon_{20}/\pi_2 > \varepsilon_{10}/\pi_1$. So, we now adjust the leisure cross-elasticities with terms representing distributional concerns. Note that distributional concerns will matter only if π_i varies across goods, which won't be the case if utility satisfies homothetic separability, i.e., has the form $u(x_0, \varphi(x_1, x_2))$, with $\varphi(\cdot)$ homogeneous in its arguments; then, consumption bundles are the same across individuals, varying only by scale.

An application is the choice of VAT rates on different commodities. We might wish to tax some goods more heavily for efficiency reasons but less heavily for equity reasons. This could help explain why existing VATs impose lower rates of tax on necessities such as food, even though necessities typically have lower own elasticities of demand (and hence in general lower cross-elasticities of demand with respect to other commodities, such as leisure). But what if we could expand our set of tax instruments a bit? The individual's budget constraint in the three-good problem considered here is $p_1(1+\theta_1)x_1 + p_2(1+\theta_2)x_2 = -x_0$, where $-x_0$ is labor income and θ_i is the proportional tax on good i. Note that we could also write this budget constraint as

$$p_1 x_1 + p_2 \frac{(1+\theta_2)}{(1+\theta_1)} x_2 = \frac{-x_0}{(1+\theta_1)}, \qquad \text{or} \qquad p_1 x_1 + p_2(1+\tau_2)x_2 = (1-\tau_0)(-x_0)$$

(Here, the tax on labor, τ_0, is expressed on a tax inclusive basis, applying to all labor income; the consumption tax is expressed on a tax exclusive basis, applying to net consumption expenditures rather than expenditures inclusive of tax. We could express either using the alternate convention, but this is typically how consumption taxes and income taxes are expressed.) That is, since the choice of the untaxed good is arbitrary, we could also have considered the problem as one with taxes on goods 0 and 2 – a labor income tax plus a separate tax on good 2. If the prior analysis had led us to choose equal taxes on commodities 1 and 2, we would now wish to tax only labor income – a labor income tax is equivalent in this model to a uniform consumption tax. Suppose that, in addition to the labor income tax and a tax on good 2, we also had available a *uniform* lump-sum tax, say T. (Note that we are not assuming that we can impose lump-sum taxes that vary across individuals.) Then, the budget constraint would involve a tax on good 2 plus a linear income tax on labor income, of the form $T + \tau_0(-x_0)$. With this additional tax instrument, when would we want to utilize the consumption tax on good 2? Not surprisingly, with an additional tax instrument, the condition is weaker than before; a sufficient condition (see Auerbach and Hines, p. 1372) is that households have separable utility with linear Engel curves with the same slopes, for which homothetic separability and equal bundles across incomes is a sufficient condition but *not* a necessary one. Indeed, allowing for a more general, nonlinear labor income tax, for which the mathematical derivation is more complex, an even weaker sufficient condition for uniform commodity taxation holds, that the utility function has the form $u(x_0, \varphi(x_1, x_2))$, i.e., is weakly separable, with no restriction at all on the shape of Engel curves (Atkinson and Stiglitz, 1976). A puzzle is why most countries with general, progressive income taxes still impose VATs with rates typically much lower (or zero) for necessities like food, or undertake other tax and regulatory policies that are strongly influenced by distributional considerations. (See, e.g., Kaplow, *National Tax Journal* 2020.)

The Production Efficiency Theorem

Let us modify the general optimal tax analysis, with heterogeneity, to allow producer prices to vary. That is, rather than assuming that the producer price vector $\boldsymbol{q}$ is fixed, assume that it is determined by efficient production behavior, and that production is determined by a constant-returns-to-scale function $f(\boldsymbol{Z}) \leq 0$, where $\boldsymbol{Z}$ is the vector of inputs and outputs. Given that relative prices may vary as we impose taxes, we express the government's revenue requirement in terms of a quantity vector of goods the government wishes to purchase, $\boldsymbol{R}$. Rather than writing down a separate government budget constraint, we may combine it with the production constraint by writing $f(\boldsymbol{X}+\boldsymbol{R}) \leq 0$, where $\boldsymbol{X}$ is, as before, the aggregate private vector of inputs and outputs.

We wish to maximize the Lagrangian, $W(V^1(\boldsymbol{p}), V^2(\boldsymbol{p}), \ldots, V^H(\boldsymbol{p})) - \mu f(\boldsymbol{X}+\boldsymbol{R})$, with respect to taxes. However, under normal circumstances (see Auerbach and Hines, footnote 15), we can maximize with respect to prices, as any vector of taxes can be achieved through a choice of prices. The first-order conditions are:

$$(4) \qquad -\sum_h W_h \lambda^h x_i^h - \mu\left[\sum_j f_j \sum_h \frac{dx_j^h}{dp_i}\right] = 0 \quad \forall i$$

Without loss of generality we can choose the units of production are such that $f_0 = 1$, and hence $f_0 = q_0$. Since production efficiency implies that $f_i/f_j = q_i/q_j \;\; \forall i, j$, it follows that $f_i = q_i \; \forall i$. Also, since for each h, $\boldsymbol{p'x}^h = 0$, it follows that $x_i^h + \sum_j p_j\, dx_j^h/dp_i = 0$. Therefore, we can subtract $x_i^h + \sum_j p_j\, dx_j^h/dp_i$ from the term in brackets in (4) to obtain:

$$(5) \qquad -\sum_h W_h \lambda^h x_i^h + \mu\left[X_i + \sum_j t_j \sum_h \frac{dx_j^h}{dp_i}\right] = 0 \quad \forall i$$

which is identical to expression (1). That is, the standard optimal tax results are not changed by the assumption that producer prices may vary, if there are no pure profits (i.e., under constant returns to scale). If there are pure profits, the result still holds, but only if the profits are first taxed away (see Auerbach and Hines, p. 1367). Intuitively, if there are constant returns to scale, producer prices may vary, but, in equilibrium, the producer of any good faces constant costs, just as in the case where prices are fixed. Thus, only demand-side terms enter into the expression.

We have assumed thus far that production is efficient. This means not only the absence of market failures on the production side, but also no government policy interventions *within* the production sector (for example, a wage subsidy for some producers but not others.) But the intuition of second-best theory suggests that we might want to use such interventions as well.

Assume now that there are two production sectors, with production functions and vectors $f(\boldsymbol{Z})$ and $g(\boldsymbol{S})$, both constant returns to scale. Also assume that production *in each sector* is efficient, but that overall production may not be. For example, we may provide subsidies to widget production in sector $g(\cdot)$ but not sector $f(\cdot)$. Let us assume the government chooses $\boldsymbol{S}$ directly, although it could accomplish this indirectly through the use of sector-specific taxes and subsidies. Then, using the fact that private plus public consumption equals total production, i.e., $\boldsymbol{X}+\boldsymbol{R} = \boldsymbol{Z}+\boldsymbol{S}$, we seek to maximize the Lagrangian

$W(V^1(\boldsymbol{p})\ V^2(\boldsymbol{p}), \ldots, V^H(\boldsymbol{p})) - \mu f(\boldsymbol{X} + \boldsymbol{R} - \boldsymbol{S}) - \zeta g(\boldsymbol{S})$

with respect to $\boldsymbol{p}$ and $\boldsymbol{S}$. The first-order conditions for $\boldsymbol{p}$ are the same as before. For $\boldsymbol{S}$, we get:

$$\mu f_i = \zeta g_i \qquad \forall i$$

which implies that the marginal rates of transformation on all margins must be the same in the two sectors, i.e., $f_i/f_j = g_i/g_j$. This is the Diamond-Mirrlees production efficiency theorem. Even though there are existing distortions, production distortions don't contribute anything (contrary to general second-best reasoning) because they effectively achieve consumption distortions indirectly (for example, raising the output price of a good whose inputs are taxed in one of the two production sectors) while *also* pushing production inside the production frontier. If we can achieve consumption distortions directly, we are better off doing so, because we will achieve an outcome that Pareto-dominates the one based on the production distortion.

Provision of Public Goods using Distortionary Taxation

Following Auerbach and Hines (pp. 1384-5), let us consider the optimal provision of a public good, G, using distortionary taxation. Assume that there are H identical individuals (heterogeneity won't add much of interest here) and that society's CRS production function is $f(\boldsymbol{X}, G) \leq 0$, where $\boldsymbol{X}$ is the vector of private consumption. The representative individual's utility function is $U(\boldsymbol{x}^h, G)$, where $\boldsymbol{X} = \sum_h \boldsymbol{x}^h$. The individual's corresponding indirect utility function may be written $V(\boldsymbol{p}; G)$, where the presence of G indicates that this is not a choice variable for individuals, but simply something that influences utility, with the property that $U_G = V_G$. Attaching the Lagrange multiplier μ to the production constraint and maximizing social welfare $H\,V(\boldsymbol{p}; G)$ with respect to the choice of prices and the level of public goods provision, we will get the same first-order conditions for $\boldsymbol{p}$ as before (since G is held constant in deriving these conditions). The first-order condition with respect to G may be rearranged as:

$$(6) \qquad H\frac{U_G}{U_0} = \frac{\mu}{\lambda}\left[\frac{f_G}{f_0} - \frac{dR}{dG}\right]$$

where good zero is the numeraire commodity (for which the tax is set equal to zero and price equal to 1), λ is the private marginal utility of income, = U_0, and dR/dG is the change in revenue resulting from an increase in public goods spending. Expression (6) includes the basic elements of the Samuelson rule (ΣMRS = MRT), but there are two modifications, the ratio μ/λ and the revenue effect dR/dG. To interpret these modifications, it is helpful to rewrite (6), using our previous definition of the *social* marginal utility of income $\alpha = \lambda + \mu \sum_j t_j \frac{dX_j}{dy} = \lambda + \mu \frac{dR}{dy}$, as

$$(6') \qquad H\frac{U_G}{U_0} = \frac{\mu(f_G/f_0) - \mu dR/dG}{\alpha - \mu dR/dy}$$

If we ignore the terms dR/dG and dR/dy, expression (6′) calls for adjusting the social cost of providing public goods, f_G/f_0, by the term $\mu/\alpha > 1$, the cost of raising funds in a distortionary manner rather than through lump-sum taxation. But, increasing public goods may provide an added benefit by causing individuals to spend more on taxed goods, raising government revenue and reducing the need for distortionary taxes – a benefit of $\mu dR/dG$ that reduces the social cost of providing public goods. On the other hand, increasing public goods spending requires increasing revenue, which reduces real income. If that real income loss reduces spending on taxed goods (i.e., $dR/dy > 0$), this raises the cost of providing public goods, by $\mu dR/dy$. As emphasized in Hendren (*Tax Policy and the Economy*, 2016), the marginal cost of public funds – the amount by which we must adjust the direct revenue cost to account for the associated deadweight loss – depends on the policy experiment. Here, the real income loss and increase in public goods spending each may interact with preexisting distortions and affect marginal deadweight loss.

It is important to keep in mind that expression (6) or (6′) indicates how the marginal condition for provision of public goods relative to a particular private good is affected. It does not tell us anything about the margins relative to other private goods, or about the *level* of public goods. Consider an example in which there are two private goods, consumption (c) and labor (L), as well as the public good; let us also assume that public good provision has no impact on revenue, i.e., $dR/dG = 0$. The individual household's budget constraint is $pc = wL$, and we can impose a consumption tax or a labor income tax, in either case letting the other good be the numeraire commodity. If we impose a consumption tax, and consumption is a normal good, then $dR/dy > 0$. Thus, $\lambda = \alpha - \mu dR/dy < \alpha < \mu$. Thus, $\mu/\lambda > 1$, so expression (6) implies that $HU_G/U_L > f_G/f_L$ – the valuation of the public good relative to labor should exceed its marginal production cost in units of labor. But suppose we impose the tax on labor, letting consumption be numeraire. If *leisure* is a normal good, then *labor* will decline with income, and so will revenue; i.e., $dR/dy < 0$. This means that $\lambda > \alpha$; in fact, as shown in Auerbach and Hines (p. 1386), $\lambda = \mu$ if preferences are Cobb-Douglas, in which case expression (6) implies that $HU_G/U_c = f_G/f_c$ – the valuation of the public good relative to consumption should equal its marginal cost in units of consumption. But, since taxing consumption and taxing labor must yield the same underlying equilibrium, these two results together imply (for Cobb-Douglas preferences) that there should be a distortion on the margin between labor and the public good, but no distortion on the margin between consumption and the public good. Put another way, there should be a distortion between goods and labor, but not between the two goods. This result may be seen as an analogy to the case with two private consumption goods and labor, where imposing a uniform tax on the two goods, or a tax on labor, distorts the labor-goods margin but not the margin between the two private goods. In both cases, the fact that there is no distortion on one margin doesn't imply that there are no distortions. In the case of public goods, we will see a reduction in the consumption of both private and public goods as we distort the labor-leisure choice.