Economics 230a, Fall 2013
Lecture Note 3: Further Optimal Tax Results

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 of 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 social marginal 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 $\frac{\mu - \alpha_i}{\alpha_i}$. Now, it is no longer optimal to set the marginal cost of funds (revenue plus excess burden per unit of revenue) equal for all revenue sources; we now wish to take into account who consumes the goods being taxed. In particular, 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. In other words, 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 – and higher taxes on goods purchased relatively intensively by those with lower social income weights. 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:

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

Since only the first terms in numerator and denominator differ, the proportional tax on good 1 will now be higher than the tax on good 2 if and only if $\frac{\varepsilon_{20}}{\pi_2} > \frac{\varepsilon_{10}}{\pi_1}$. 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. Here, 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 <u>tax inclusive basis</u>, applying to all labor income; the consumption tax is expressed on a <u>tax exclusive basis</u>, 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 <u>linear income tax</u> 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 *not* a necessary condition. Indeed, if we allowed 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 would hold, that the utility function is weakly separable, i.e., has the form $u(x_0, \varphi(x_1, x_2))$, with no restriction at all on the shape of Engel curves (Atkinson and Stiglitz, 1976).

Application: Tax Treatment of the Family

A classic application of optimal tax theory is the treatment of the family. Let the three goods, x_0, x_1, and x_2 now be consumption, husband's labor, and wife's labor, respectively, and let good 0 (consumption) be the untaxed numeraire commodity. Assuming that elasticities ε_{12} and ε_{21} are zero (or, more generally, small), we can apply the inverse elasticity rule when only efficiency concerns matter, and tax more heavily the income of the spouse with the lower compensated labor supply elasticity; empirical evidence suggests that this would be the husband. A second step to consider, though, is distributional concerns, where issues like the extent of assortative mating come up. For example, will women with high incomes typically be found in families with high incomes? A third issue that is relevant is how families make decisions. The standard optimal tax approach treats the family as a single optimizing unit, but given empirical evidence other approaches may be more plausible, such as intrafamily bargaining. The paper by Alesina et al. considers the optimal taxation of a representative couple (so there is no issue of distribution across families), but it assumes intrafamily bargaining and also generates differing labor supply elasticities endogenously, as a consequence of differences in bargaining power or comparative advantage. A key result of the paper is that, if government can make transfers *within* the family (which matter given that bargaining determines outcomes), then the standard result that men should face higher marginal tax rates than women still generally holds.

One further issue that the paper (and the previous literature) generally ignores is that there are couples as well as single taxpayers. How to tax on singles vs. couples is a complicated question, not only because the marriage decision may be affected, but also because it is not obvious how to compare one-individual and two-individual units. The US has separate tax schedules for single individuals and married couples, while many other countries use one schedule for individual taxation, regardless of whether the individual is married. Even in such countries, though, elements of the transfer system, such as low-income payments, are often family based. This is how the UK tax and transfer system operates, for example.

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:

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

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, $\boldsymbol{p}'\boldsymbol{x}^h = 0$, it follows that $x_i^h + \sum_j p_j \frac{dx_j^h}{dp_i} = 0$.

Therefore, we can subtract $x_i^h + \sum_j p_j \frac{dx_j^h}{dp_i}$ from the term in brackets in (4) to obtain:

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

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 optimal tax expression. This does not mean that the *equilibrium* will be the same in the two cases, since changes in the vector $\boldsymbol{X}$ will depend on demand and supply responses.

We have assumed thus far that production is efficient. A sufficient condition for this to be true is that there are no market failures *and* no government policy interventions *within* the production sector. (One example of such an intervention would be an industry-specific wage subsidy.) But the intuition of second-best theory suggests that we might want to use such interventions as well, given that we are not at a Pareto optimum.

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.

Applications: social discount rate; tariffs and export subsidies.