Lecture 5

Outline

1. Many-person Ramsey Rule
2. Commodity taxes and income taxes
3. Indirect taxes and direct taxes
4. Kaplow paper

1. Many-Person Ramsey Rule
   (a) The derivation is messy. What it basically does is make the equity/efficiency tradeoff precise.
      i. Taxes have direct and indirect effects. Both affect the final policy.
      ii. People respond to taxes. This affects their utility and the revenue the government raises.
      The effects on individual utility influence policy directly if these utilities matter a lot in the social welfare function.
      They influence policy indirectly, however, through the revenue effects. Even people favored by the social welfare function will have to be taxed a lot if others sharply reduce their spending on goods they alone consume in response to taxes.
      iii. One could say that the purpose of the theory is to give a full accounting of these effects.
   (b) $i = 0, ..., N$, goods.
   $h = 1, ..., H$, people.
   x_i^h, "demand by person h for good i"
   We assume linear technology. Recall, this lets us drop all of the market clearing equations and leaves us with just the government's revenue constraint.
   Note of notation! Person h pays tax on good k equal to $t_k x_k^h$. The effect of this tax on her demand for good i is $\frac{\partial x_i^h}{\partial q_k}$, given linear technology.
   (c) Optimization:

$$\begin{array}{ll} \underset{(t_1, ..., t_n)}{\text{Max}} & W[V^1(p^* + t), ..., V^H(p^* + t)] \\ \text{s.t.} & \sum_{h=1}^{H} \sum_{i=1}^{n} t_i x_i^h = R \\ & p_0^* = 1, t_0 = 0, h = 1, ..., H \end{array}$$

Note! As before, I am deleting any reference to income from indirect utility since there is no profit income. This becomes a slight problem when we want to talk about income effects – derivatives with respect to income. It should be understood that indirect utility is still a function of income, it is just evaluated at zero income, and we can still take derivatives with respect to income, also evaluated at zero income.

(d)
$$\mathcal{L} = W[V^1(p^* + t), ..., V^H(p^* + t)] + \lambda \left[\sum_{h=1}^{H} \sum_{i=1}^{n} t_i x_i^h - R \right]$$

(e) Consider the first order condition for the kth tax.

$$\frac{\partial \mathcal{L}}{\partial t_k} = \left(\frac{\partial W}{\partial V^1} \frac{\partial V^1}{\partial q_k} + ... + \frac{\partial W}{\partial V^H} \frac{\partial V^H}{\partial q_k} \right) + \lambda \sum_{h=1}^{H} \left(x_k^h + \sum_{i=1}^{n} t_i \frac{\partial x_i^h}{\partial q_k} \right) = 0$$

(f) Using Roy's identity, we have:

$$\frac{\partial W}{\partial V^h} \frac{\partial V^h}{\partial q_k} = \frac{\partial W}{\partial V^h} (-\alpha^h) x_k^h$$

where α^h is, as always, the Lagrange multiplier from h's utility maximization problem.

Define:

$$\beta^h \equiv \frac{\partial W}{\partial V^h} \alpha^h > 0$$

This is the *social marginal utility of income for person h* (also called h's social marginal utility of consumption): the social evaluation of the increase in utility of person h made possible when h is endowed with an extra unit of numeraire.

Therefore:

$$\frac{\partial W}{\partial V^h} \frac{\partial V^h}{\partial q_k} = -\beta^h x_k^h$$

(g) Recall the Slutsky equation:

$$\frac{\partial x_i^h}{\partial q_k} = S_{ik}^h - x_k^h \frac{\partial x_i^h}{\partial I^h}$$

(h) Make both substitutions into the first order condition for the kth tax:

$$-\sum_{h=1}^{H} \beta^h x_k^h + \lambda \sum_{h=1}^{H} x_k^h + \lambda \sum_{h=1}^{H} \sum_{i=1}^{n} t_i S_{ik}^h - \lambda \sum_{h=1}^{H} \sum_{i=1}^{n} t_i x_k^h \frac{\partial x_i^h}{\partial I^h} = 0$$

Divide every term by $\sum_{h=1}^{H} x_k^h$:

$$-\frac{\sum_{h=1}^{H} \beta^h x_k^h}{\sum_{h=1}^{H} x_k^h} + \lambda + \lambda \frac{\sum_{h=1}^{H} \sum_{i=1}^{n} t_i S_{ik}^h}{\sum_{h=1}^{H} x_k^h} - \lambda \frac{\sum_{h=1}^{H} \sum_{i=1}^{n} t_i x_k^h \frac{\partial x_i^h}{\partial I^h}}{\sum_{h=1}^{H} x_k^h} = 0$$

Divide by λ and rearrange:

$$\frac{\sum_{h=1}^{H}\sum_{i=1}^{n} t_i S_{ik}^h}{\sum_{h=1}^{H} x_k^h} = \frac{1}{\lambda}\frac{\sum_{h=1}^{H}\beta^h x_k^h}{\sum_{h=1}^{H} x_k^h} + \frac{\sum_{h=1}^{H}\left(\sum_{i=1}^{n} t_i \frac{\partial x_i^h}{\partial I^h}\right) x_k^h}{\sum_{h=1}^{H} x_k^h} - 1 \qquad (1)$$

$k = 1, ..., n$

(i) We can interpret the left hand term as in the case of one individual. If we define the aggregate compensated demand for good k at the optimum:

$$X_k^c(p^* + t, U^{1*}, ..., U^{n*}) \equiv \sum_{h=1}^{H} x_k^{h,c}(p^* + t, U^{h*}) = \sum_{h=1}^{H} x_k^h(p^* + t)$$

(since $x_k^{h,c}(p^* + t, U^{h*}) = x_k^h(p^* + t)$). Then:

$$\frac{X_k^c(p^* + t, .) - X_k^c(p^*, .)}{X_k^c(p^* + t, .)} = \frac{\sum_{h=1}^{H}\sum_{i=1}^{n} t_i S_{ik}^h}{\sum_{h=1}^{H} x_k^h}$$

where we use the assumption that compensated demand is linear in the relevant range. This says that the term on the right is the (negative) percentage change in compensated demand from eliminating all taxes.

This is usually interpreted as the *reduction* in compensated demand from *imposing* the optimal tax vector. This is not quite right.

(j) Unlike before, the right hand side of (1) is not independent of good k. So, we do not obtain a simple "equal percentage change" rule.

Furthermore, we can not sign the right hand side as we could before. It is reasonable to suppose it is negative, so eliminating all taxes increases aggregate compensated demand, but this is not a direct implication of negative definiteness of the Slutsky matrix.

Nevertheless, the equation gives a little intuition about the properties of the optimal tax system.

Assume that the right hand side is negative. Also, we will follow standard usage and refer to "reduction" in compensated demand.

i. The reduction in compensated demand for good k should be small (in absolute value) if $\sum_{h=1}^{H}\beta^h x_k^h$ is large.
This occurs if good k is largely demanded by people whose social marginal valuation of income is large, which means they matter a lot in the social welfare function, and whose personal marginal utility of income is large (they are poor).

ii. The reduction in compensated demand for good k should be small if $\sum_{h=1}^{H}\left(\sum_{i=1}^{n} t_i \frac{\partial x_i^h}{\partial I^h}\right) x_k^h$ is large. This occurs if good k is largely demanded by people who sharply reduce their demands for taxed goods when their income changes. Taxing good k is inefficient since the tax rate would have to be very high to meet the budget constraint, leading to more distortion in the aggregate than is necessary. Less distortion results from taxing other goods at more moderate rates.

(k) A little further insight into the many-person Ramsey rule comes from considering when it reduces to the one-person Ramsey rule. That is to say, from considering when the equity component of optimal taxes disappears and only the efficiency component remains.

Define:

$$\bar{x}_k = \frac{\sum_{h=1}^{H} x_k^h}{H}$$

and:

$$b^h = \frac{\beta^h}{\lambda} + \sum_{i=1}^{n} t_i \frac{\partial x_i^h}{\partial I^h}$$

Then making the substitutions gives:

$$\begin{aligned} \frac{\sum_{h=1}^{H} \sum_{i=1}^{n} t_i S_{ik}^h}{\sum_{h=1}^{H} x_k^h} &= \sum_{h=1}^{H} \frac{x_k^h}{H\bar{x}_k} \left[\frac{\beta^h}{\lambda} + \sum_{i=1}^{n} t_i \frac{\partial x_i^h}{\partial I^h} \right] - 1 \\ &= \sum_{h=1}^{H} \frac{b^h x_k^h}{H\bar{x}_k} - 1, \quad k = 1, ..., n \end{aligned}$$

This reduces to the one-person Ramsey rule in two cases:

i. $$b^h = b, \quad h = 1, ..., H$$

In this case the right hand side reduces to $b - 1$, which is independent of h and k.

This says that all individuals are valued equally and have the same propensity to pay taxes. In this case it is unnecessary to use commodity taxes to subsidize some people at the expense of others.

ii. $$\frac{x_k^h}{\bar{x}_k} = \hat{x}^h, \quad k = 1, ..., n$$

Again, the right hand side is independent of h and k.

This says that individual h consumes a common fraction of each good. The fraction may vary across individuals but the pattern of expenditure is the same across individuals. For example, the percentage of each good consumed by each would look something like:

	Person		
Good	1	2	3
1	20%	50%	30%
2	20%	50%	30%
3	20%	50%	30%
4	20%	50%	30%
⋮	⋮	⋮	⋮

In this case it is infeasible to use commodity taxes to subsidize some people at the expense of others. Revenue raised from any person via a tax on, say, good 2 would just give it back to him via a balanced-budget subsidy on good 3.

(l) There is a third set of results that comes from redoing the optimization assuming the government can also take a fixed amount of income from all individuals (a "poll tax").

See Auerbach or Diamond (1975) for details.

2. Commodity taxes and income taxes

(a) The fundamental questions here are, when do we need the income tax if we have a full set of commodity taxes; and when do we need any commodity taxes if we have the income tax.

(b) There are a variety of frameworks of analysis. The key variations are:

i. One person versus many people.

ii. The functional form of the tax schedule. Is it an anonymous and proportional function of income, so taxes are just $\tau w^h L^h$ for person h? Or can it be any function of income, $T(w^h L^h)$?

iii. The information known to the government when choosing the tax schedule. Does it observe individual characteristics and prices, including the wages each person earns?

If the tax is proportional, then it would be a function of this information, something like $\tau(.)w^h L^h$ for person h. If it can be any function of this information, then $T(., w^h L^h)$.

(c) Consider the one-person optimal commodity tax model.

Consider the question, when do we need the income tax if we have a full set of commodity taxes?

Well, we know that there is no loss of generality in leaving one good or factor untaxed. The government can therefore leave labor supply untaxed without any loss.

So, the answer in this simple case is, "never."

We get the same answer in the many-person optimal commodity tax model.

(d) Now consider the one-person optimal commodity tax model and the converse question, when do we need any commodity taxes if we have the income tax?

An obvious answer is, if the solution to the OT problem is a uniform commodity tax, then we do not need any commodity taxes. When this occurs we can replace the uniform commodity tax with a proportional income tax. This is just accounting.

Proof:

i. We have been writing individual h's budget constraint as:

$$\sum_{i=0}^{n} q_i x_i^h = 0$$

With the notation $x_0^h = -L^h$ and $q_0 = w$, this becomes:

$$wL^h = \sum_{i=1}^{n} q_i x_i^h$$

Now impose a proportional income tax at rate τ:

$$w(1-\tau)L^h = \sum_{i=1}^{N} q_i x_i^h$$

Dividing both sides by $1-\tau$ gives the equivalent equation:

$$wL^h = \sum_{i=1}^{N} \frac{q_i}{1-\tau} x_i^h$$

Thus, there is a uniform proportional tax on all commodities that leaves the individual with the same budget constraint that she had with the income tax. The uniform proportional tax on all commodities could be achieved through an appropriate system of unit taxes on the commodities. Similarly, the proportional income tax could be achieved through the appropriate unit tax on labor supply.

So, when is a uniform proportional tax on all commodities optimal?

Myles (125-127) gives a nice discussion. There is also a recent paper on this topic (Besley and Jewitt (1990), *Econometrica*).

The *necessary* conditions for this to hold are very strong: they are somewhat like homotheticity conditions. So, as Myles says, "there is no reason they should be satisfied in practice."

(e) Atkinson and Stiglitz (section 14-3) go to the opposite extreme. They use the optimal income tax framework. This has multiple individuals, and the government chooses an income tax schedule that is a function of total labor income only.

More specifically:

i. Individuals are assumed identical except in the wage they receive. Preferences are assumed weakly separable in labor. This means utility can be written:

$$U[L, C(X_1, ..., X_n)]$$

ii. The government can levy a non-linear income tax. So, τ is a function of $-wL$.

However, the tax cannot take into account any individual characteristics (it is "undifferentiated").

iii. The claim is that commodity taxes are not necessary.

Given separability and identical preferences, variation in consumption is closely tied to variation in income.

As long as the income tax rate can vary with income, this correlation makes consumption taxes redundant. They provide no extra capacity to meet equity and efficiency goals.

iv. The meaning of this result depends on whether you think weak separability is a weak or strong assumption.

v. There is also an interpretation in terms of the theory of the second best (Atkinson-Stiglitz). We have to create distortions to raise revenue. In general, there is no presumption that we want to approximate the first best result (lump sum taxation) by keeping the number of markets in which we create distortions to a minimum. Nevertheless, if the separability assumption holds, then we do keep the number of markets to a minimum by just intervening in the labor market.

3. Indirect taxes and direct taxes

(a) In standard usage, "direct" tax means "income tax" and "indirect tax" means "commodity tax."

(b) In the optimal commodity tax model, however, this distinction does not exist. Labor is a commodity and there is nothing special about "income" as a tax base. One supplies a factor and receives money instead of demanding a good and paying money.

The distinction between "direct" and "indirect" is based on differences between income and commodity taxes that are not part of this model.

(c) In reality, the income tax is not administered like a commodity tax (at point of sale). As a practical matter, this difference permits one's income tax liability to be highly "personalized":

i. It may be non-linear in income.

ii. It may depend on how one spends one's income (for example through tax deductions).

iii. It may depend on characteristics of the individual or household.

iv. On the other hand, since they are not levied at the "point of sale" (the firm), it is assumed that the government does not know how many hours an individual works. Thus, the government cannot levy a unit tax on hours worked. However, we know from the previous analysis that *this* kind of income tax is precisely the kind that is *not* needed when commodity taxes are available.

(d) In contrast, liability from commodity taxes is a simple function of a few variables.

i. It is linear in the quantity purchased (Atkinson and Stiglitz consider the possibility of non-linear commodity taxes, but this is an unusual case).

ii. It does not depend on the quantities of other goods purchased.

iii. It does not depend on the purchaser's income.

iv. It does not depend on characteristics of the purchaser or the purchaser's household.

(e) These considerations lead Atkinson and Stiglitz to define a *direct tax* as one that may be adjusted to the individual characteristics of the taxpayer. Thus, the income tax as we have it is a direct tax, but the simple wage tax (that need not be used when a full set of commodity taxes are available) is not.

In contrast, an *indirect tax* is levied on the transactions irrespective of the circumstances of the buyer or seller. Commodity taxes as we have been discussing them are indirect taxes.

(f) The U.S. Constitution refers to direct taxes in Article 1 Section 9 and sharply restricts their use. Any tax in this category must raise revenue in proportion to the population of each state.

This is known as the, "apportionment rule."

Formally:

$$\frac{\text{Pop}_A}{\text{Pop}_B} = \frac{\text{Total Revenue}_A}{\text{Total Revenue}_B}$$

So, one obvious implication is that more populous states must pay more. However, there is a less obvious implication.

$$\frac{\text{Pop}_A}{\text{Pop}_B} = \frac{\text{Total Revenue}_A}{\text{Total Revenue}_B} = \frac{t_A \text{Pop}_A \text{Tax Base-per-Capita}_A}{t_B \text{Pop}_B \text{Tax Base-per-Capita}_B}$$

So

$$\frac{t_A}{t_B} = \frac{\text{Tax Base-Per-Capita}_B}{\text{Tax Base-Per-Capita}_A}$$

Insofar as *wealthy* states have higher tax-base-per-capita, the rule says that wealthier states must have lower tax rates. If B is a wealthy state, so it has twice the base of state A, then the tax rate there must be *half* as large as in state A.

This was adopted as part of a bargain with the small wealthy states, giving them some tax protection in return for their willingness to accept a smaller number of representatives in Congress.

(g) Interestingly, when the Federal income tax was first proposed, there was substantial disagreement over whether it was a direct tax within the meaning of the Constitution. Eventually the Supreme Court decided it was. This led to the Sixteenth Amendment (1913), which says that apportionment is not necessary.

4. Kaplow, "On the undesireability of commodity taxation even when income taxation is not optimal," Journal of Public Economics 90 (2006).

   There are problems with this paper!

   Suppose we have two commodities and labor, so we have $U[v(x_1, x_2), L]$. Commodity and income taxes are arbitrary and the person maximizes utility subject to this constraint. This gives the initial equilibrium, x_1', x_2', L'.

   A property of the initial equilibrium is that, if you maximize the subutility function $v(x_1, x_2)$ subject to the original constraint and L' given, you would again get x_1' and x_2'. Actually, this has nothing to do with separability, it is an obvious property of optima characterized by sets of first-order conditions. It is clearly true. So the indifference curve $v(x_1, x_2) = v(x_1', x_2')$ is tangent to the original constraint and L' given, i.e., in (x_1, x_2) space.

   Now remove the commodity taxes and adjust the tax rate on labor so the new constraint is tangent again to $v(x_1, x_2) = v(x_1', x_2')$. Specifically, assuming labor income (wL) does not change, levy a lump-sum tax on labor income. The government can do this because it has complete information about preferences (the utility function). Call the new tangency (x_1^*, x_2^*). Kaplow's claims are that if you now repeat the problem of maximizing (x_1, x_2, L) subject to this new budget constraint you obtain (x_1^*, x_2^*, L') and raise strictly more revenue. Note that the government never knows or observes L', but it does know and observe that labor income does not change. So far, so good.

   The problem comes when you try to do this for a group of individuals. What he must be saying is that the government computes a set of lump-sum modifications to the original tax schedule, one for each individual and dependent on the initial observed choices. You can do the construction, but now present the new and complete tax schedule to the group and say "optimize." What is going to happen? We know that no will want to make a local adjustment, but they may want to make a global adjustment – mimic other types – if the incentive compatibility constraint doesn't bind. And, I see no reason why it should necessarily bind.

The challenge now is to create an example. Use two types. Construct two schedules, one for each individual; write down the composite schedule from the separate schedules; present the composite to the two individuals; show that one changes labor supply. This will falsify his Lemma 1.