Lectures Notes on Economics of Taxation

Hanming Fang

October 14, 2004

Contents

1 Tax Incidence

1.1 Partial Equilibrium: The Competitive Case

Let $L^d(w)$ and $L^s(w)$ denote the labor demand and labor supply respectively where w is the *net* wage. Suppose that a payroll tax at rate t is introduced. The equilibrium is given by

$$L^d(w(1+t)) = L^s(w).$$

Start from $t = 0$ and differentiate w.r.t t we obtain

$$\begin{aligned} L^{d\prime}(dw(1+t) + wdt)\Big|_{t=0} &= L^{s\prime}(w)\,dw \\ L^{d\prime}(dw + wdt) &= L^{s\prime}(w)\,dw \end{aligned}$$

Divide both sides by wdt, we have

$$L^{d\prime}\left(\frac{dw}{wdt} + 1\right) = L^{s\prime}(w)\frac{dw}{wdt}$$

Note that

$$\frac{dw}{wdt} = \frac{d\ln w}{dt}$$

we have

$$\begin{aligned} \frac{d\ln w}{dt} &= \frac{L^{d\prime}}{L^{s\prime} - L^{d\prime}} \\ &= \frac{L^{d\prime}w/L^d}{L^{s\prime}w/L^s - L^{d\prime}w/L^d} \\ &= -\frac{\varepsilon_D}{\varepsilon_S + \varepsilon_D} \end{aligned}$$

where

$$\varepsilon_D = -\frac{wL^{d\prime}}{L}, \quad \varepsilon_S = \frac{wL^{s\prime}}{L}$$

are the wage elasticity of labor demand and supply respectively.

Let $W = w(1+t)$ be the gross wage. Then

$$\begin{aligned}\frac{\partial \ln W}{\partial t} &= \frac{\partial W}{W \partial t} = \left.\frac{w + \frac{\partial w}{\partial t}(1+t)}{w(1+t)}\right|_{t=0} \\ &= 1 + \frac{d \ln w}{dt} \\ &= \frac{\varepsilon_S}{\varepsilon_S + \varepsilon_D}\end{aligned}$$

Finally, the fall in employment is given by

$$-\frac{\partial \ln L^S}{\partial t} = -\frac{L^{S\prime} \partial w}{L^S \partial t} = -\frac{\varepsilon_S}{w}\frac{\partial w}{\partial t} = \varepsilon_S \frac{d \ln w}{dt} = \frac{\varepsilon_S \varepsilon_D}{\varepsilon_S + \varepsilon_D}$$

Alternatively, the fall in employment can be calculated as

$$-\frac{\partial \ln L^D}{\partial t} = -\frac{L^{D\prime} \partial W}{L^D \partial t} = \frac{\varepsilon_D}{W}\frac{\partial W}{\partial t} = \varepsilon_D \frac{d \ln W}{dt} = \frac{\varepsilon_S \varepsilon_D}{\varepsilon_S + \varepsilon_D}$$

- Net wage decreases all the more that demand is more elastic relative to supply;
- Gross wage increases all the more that demand is less elastic relative to supply;
- The fall in employment is all the larger that demand and supply are more elastic.

1.2 Partial Equilibrium: The Monopoly Case

Let p denote the price that consumer pays. Let t be the tax rate. Thus the producer price is $p/(1+t)$. Let a monopolist firm solve

$$\max_{\{p\}} \left\{ \frac{p}{1+t} D(p) - C(D(p)) \right\}$$

The first order condition is

$$\begin{aligned}\frac{D(p)}{1+t} + \frac{p}{1+t} D'(p) &= C'(D(p)) D'(p) \\ &\Rightarrow \frac{pD(p)}{p(1+t)} + \frac{p}{1+t} D'(p) = C'(D(p)) D'(p) \\ &\Rightarrow \frac{p}{1+t} = \frac{C'(D(p)) D'(p)}{\frac{D(p)}{p} + D'(p)} \\ &\Rightarrow \frac{p}{1+t} = \frac{C'(D(p))}{\frac{D(p)}{D'(p)p} + 1} = \frac{C'(D(p))}{1 - (1/\varepsilon_D(p))}.\end{aligned}$$

Now consier some special cases.

Let C' be a constant, $C' = c$. This implies that supply is perfectly elastic.

If the demand curve also have constant elasticity, i.e., $\varepsilon_D(p) = \varepsilon_D$ does not depend on p. Then indeed the above formula says that the monopolist seller will choose p such that

$$\begin{aligned} \frac{p}{1+t} &= \frac{c}{1-(1/\varepsilon_D)} \\ p &= \frac{c(1+t)}{1-(1/\varepsilon_D)}. \end{aligned}$$

The seller keeps net price constant. But the seller's profit is also decreased as t increase because the gross price increases with t.

If the demand curve is linear, say $D(p) = d - p$, then

$$\varepsilon_D(p) = \frac{p}{d-p}$$

hence

$$\frac{p}{1+t} = \frac{c}{1-\frac{d-p}{p}} = \frac{pc}{2p-d}$$

or

$$\begin{aligned} p &= \frac{c(1+t)+d}{2} \\ \frac{p}{1+t} &= \frac{c}{2} + \frac{d}{2(1+t)} \end{aligned}$$

Thus the consumer price p increases with t, and the producer price $p/(1+t)$ decreases with t. That is, both sides of the market bear some of the burden of the tax.

1.3 Specific Tax versus Ad Valorem Tax

- A specific tax is levied according to unit of commodities, it is an absolute amount. E.g. \$.18 per gallon of gas (regardless of the actual price of a gallon);

- Ad Valorem tax is proportional to the value of the unit, it is a percentage tax rate. E.g. 6% sales tax on a \$400 sofa is \$24. If the sofa price is reduced to \$200, the tax will be only \$12.

Result: Whether a tax is levied as a specific tax or an ad valorem tax does not affect allocation or incidence in competitive market as long as the *actual tax amount* is the same under the two tax forms. This is not true if the seller is a monopoly.

To see this, let t be the ad valorem tax rate. In competitive market, the market clearning condition implies that

$$D\left(p^*\left(1+t\right)\right)=S\left(p^*\right).$$

That is, $S\left(p\right)$ units will be sold, and a total tax revenue of $p^*tS\left(p^*\right)$ will be collected. If the government switches to a revenue neutral specific tax, it can simply charge a tax of $\tau=p^*t$ per unit of sale. The demand will be $D\left(p+p^*t\right)$. Note that the market will again be cleared at p^*, because p^* is a solution to $D\left(p+p^*t\right)=S\left(p\right).$

Now consider the monopoly case. Let $P\left(q\right)$ be the inverse demand function, and let $C\left(q\right)$ be the cost function. In the no-taxation case, the monopoly's optimum satisfies

$$MR\left(q\right)=C'\left(q\right)$$

where

$$MR\left(q\right)=P\left(q\right)+qP'\left(q\right)$$

is the marginal revenue.

If the monopoly pays an ad valorem tax at rate t, then the marginal revenue decreases by $tMR\left(q\right)$. A specific tax τ will reduce the marginal revenue by τ.[1] Fix a quantity q. Assume the tax parameters t and τ have been chosen so as to collect the same amount of tax revenues at production level q. Then

$$tqP\left(q\right)=\tau q,\text{ or }\tau=tP\left(q\right).$$

This implies that $\tau>tMR\left(q\right)$ because $MR\left(q\right)<P\left(q\right)$. Thus, at given production and tax revenue, the specific tax reduces marginal revenue more than the ad valorem tax. For a given tax revenue, an ad valorem tax reduces production less, which is good for social welfare since the monopoly already produces too little. Thus ad valorem taxes like VAT should be preferred to specific taxes such as some excises.

1.4 General Equilibrium

The analysis so far neglects the effects of taxes on the general price level, other than the price of the good that is being taxed, and thus neglects the substitutions among the goods

[1] To see this, note that under an ad valorem tax, the revenue is $P\left(q\right)\left(1-t\right)q$; and the revenue under a specific tax is $\left[P\left(q\right)-\tau\right]q$.

as a result. The founding model in the general equilibrium theory of tax incidence is that of Harberger (1962).

Arnold Harberger (1962). "The Incidence of the Corporation Tax." *Journal of Political Economy*, 215-40.

Any analysis of the general equilibrium effects of taxes with any attempt of realism must adopt the computable general equilibrium approach (CGE) developed after Shoven and Whalley (1972): "A General Equilibrium Calculation of the Effects of Differential Taxation of Income from Capital in the U.S." *Journal of Public Economics*: 281-321.

In general equilibrium, factors must be paid the same net-of-tax rate in both sectors, since they are perfectly mobile. This implies that, if capital taxation increases in sector 1, then the net return of capital must decrease in the whole economy, and not only for capital used in sector 1.

2 Distortions and Welfare Losses

Taxes will create distortions in an economy when its introduction changes the relative prices perceived by various agents. For example, consumer perceive after-tax price, while producers perceive before-tax prices. In such situations, equilibrium of the eocnomy does not lead to the equality of marginal rates of substiution, thus violating the conditions for Pareto optimality. The competitive market does not coordinate the agents efficiently because it sends different signals to different agents.

2.1 An Example

- Consider a two-good, one-consumer and one-firm economy;
- The consumer's utility function is

$$u(x_1, x_2) = x_1 x_2;$$

- The firm can transform good 1 into good 2 through a linear production function $X_2 = X_1/c$; Let good 1 be the numeraire;
- The consumer has one unit of endowment of good 1.

Without taxation, the equilibrium is computed as follows. The price of good 2 will be $p_2 = c$. The consumptions of good 1 and 2 are

$$x_1 = \frac{1}{2}, x_2 = \frac{1}{2c}$$

and the consumer achives a utility of

$$u = \frac{1}{4c}.$$

Suppose now we introduce a specific tax t on good 2, and then the tax revenue is distributed to the consumer as a lump-sum transfer T. In this case, the equilibrium consumer price for good 2 will be

$$p_2' = c + t.$$

Thus the optimal consumptions for good 1 and 2 are

$$x_1' = \frac{1+T}{2}, x_2' = \frac{1+T}{2(c+t)}.$$

But in equilibrium, $T = tx_2'$, which implies that

$$T = \frac{t}{2c+t}.$$

Plugging T into above, we obtain

$$\begin{aligned} x_1' &= \frac{c+t}{2c+t} \\ x_2' &= \frac{1}{2c+t}, \end{aligned}$$

and the consumer's utility is

$$u' = \frac{c+t}{(2c+t)^2}.$$

Note that

$$u - u' = \frac{t^2}{4c(2c+t)^2} > 0.$$

This is called the *deadweight loss* or *excess burden* of the tax. Note that this excess burden exists even though the government returns all the tax proceeds of the tax to the consumers.

2.2 Effects of Taxation

2.2.1 Effects of Income Tax on Labor Supply

Consumer a consumer with utility function $u(C,L)$ where C is the consumption of a numeraire good, and L is the labor supply. u is increasing in C and decreasing in L. Suppose that the wage rate is w, and the consumer also receives and unearned income R.

Suppose that a proportional income tax at rate t is imposed. Then the consumer's budget constraint is

$$C \leq (1-t)(wL+R) = sL + M$$

where $s = (1-t)w$ and $M = (1-t)R$. The consumer solves

$$\begin{aligned} & \max_{\{C,L\}} u(C,L) \\ s.t.\ C \ &\leq\ (1-t)(wL+R) \end{aligned}$$

Let $L^*(t)$ be the labor supply when the tax rate is t.

$$\frac{\partial L^*(t)}{\partial t} = \frac{\partial L^*}{\partial s}\frac{\partial s}{\partial t} + \frac{\partial L^*}{\partial M}\frac{\partial M}{\partial t}$$

Using Slutsky equation, we have

$$\frac{\partial L^*}{\partial s} = \frac{\partial \hat{L}}{\partial s} + L^*\frac{\partial L^*}{\partial M}$$

Thus

$$\frac{\partial L^*(t)}{\partial t} = -w\frac{\partial \hat{L}}{\partial s} - (wL^* + R)\frac{\partial L^*}{\partial M}$$

The first term is the substitution effect and is clearly negative; the second term are the two sources of the income effect (a lower after tax wage income and a lower after-tax unearned income). Whether the second term is positive or negative depends on whether leisure is normal good. One thing is clear, however, that the income effect is smaller for lower for low-income individuals. Thus income tax may have more disincentive effects on the poor than on the rich, other things equal.

Exercise: If $u(C,L) = a\ln C + (1-a)\ln\left(\bar{L} - L\right)$, what is the effect of tax on labor supply? [Answer: the income effect and substitution effect exactly cancels out]

2.2.2 Effects of Taxation on Savings

Here we focus on the effect of taxation of income from saving on the time profile of consumption over the life cycle, neglecting the taxation of labor income.

Consider a consumer who lives for two periods and who supplies one unit of labor inelasticly. Het works in the first period and receives a wage w, consumes some part of it and saves the rest with an interest rate r. In the second period, the agents does not work, and her interest income from earlier savings is taxed at rate t.

Suppose that the consumer's utility function is $u(C_1, C_2)$. The consumer's problem is

$$\begin{aligned} & \max u(C_1, C_2) \\ \text{s.t.} \quad & C_1 + S = w \\ & C_2 = [1 + r(1-t)] S \end{aligned}$$

Or equivalently, we can write the budget constraint of the individual as

$$C_1 + \frac{1}{1 + r(1-t)} C_2 = w.$$

Thus

$$p = \frac{1}{1 + r(1-t)}$$

is the relative price of period-2 consumption. An increase in t will lead to a higher p, making period 2 consumption more expensive.

As usual, an increase in p due to taxation has an income effect (reducing consumption in both periods) and a substitution effect (leading to a reduction of period 2 consumption and less saving).

2.3 Welfare Losses

2.3.1 Measures of Welfare: Equivalent and Compensating Variations

In economics, there is often a need for calculating welfare changes among agents when prices are changed.

- For firms, there is no real difficulty in the measure of the effects of price changes: if $\pi(p)$ is a firm's profit when the prices are p, i.e. $\pi(p) = \max_{y \in Y} p \cdot y$, then the welfare of the firm changes from $\pi(p)$ to $\pi(p')$ when the prices change from p to p'.

- For consumers, things are much more difficult. It might seem natural to say that the welfare of a consumer is given by his indirect utility,

  $$V(p, R) = \max u(x) \text{ s.t. } p \cdot x \leq R.$$

  And say that a consumer's welfare changes from $V(p, R)$ to $V(p', R)$ when the prices change from p to p'.

  The problem with this measure is that it depends on the choice of the utility function u and so it is not satisfactory.

- In 1940s, Hicks introduced two measures, called **equivalent variation** and **compensating variation.**

Equivalent Variation: The equivalent variation of income E, is the sum that must be given to (or deducted from) the consumer in the *initial sate* in order for him or her to have the same utility as in the final state. That is

$$V\left(p', R\right) = V(p, R + E)$$

Compensating Variation. The compensation variation of income C, is the sum that must be deducted from (or be given to if it is a negative deduction) the consumer in the *final state* in order for him or her to have the same utility as in the initial state. That is,

$$V\left(p', R - C\right) = V(p, R).$$

Note that both measures are quite independent of the choice of the utility function u. That is, an increasing transformation of u will leave the two measures unaffected. The reason is that the indirect utility function itself is $V(p, R) = u(x(p, R))$ where $x(p, R)$ is the Marshallian demand function.

Using the expenditure function, defined as below:

$$\begin{aligned} e(p, u) &= \min p \cdot x \\ \text{s.t. } u(x) &\geq u \end{aligned}$$

we can write

$$\begin{aligned} E &= e\left(p, V\left(p', R\right)\right) - R \\ C &= R - e\left(p', V(p, R)\right). \end{aligned}$$

Typically, these two measures give different answers.

Exercise: Show that, if the utility function u is quasilinear, i.e., there exists a good, say, the last good x_K, such that

$$u(x) = \tilde{u}(x_1, ..., x_{K-1}) + x_K,$$

then the equivalent variation and the compensating variation are the same.

2.3.2 Welfare Loss from Taxations

This is a very difficult problem. But we can get some intuition (of limited use, of course) from considering a very simple economy.

Consider an economy with I agents, and let $i = 1, ..., I$. Agent $i's$ indirect utility function is $V(p, R)$.

Suppose that at a zero tax rate $t_0 = 0$, the market clearing price vector is p_0. Now introduce a tax $t_1 = t$ on some good. Suppose that the after-tax prices move from p_0 to p_1. Further assume that the consumer's income R is unchanged by the taxes.

If we use the equivalent valuation to measure the change of the consumer welfare, we know that the tax creates a welfare loss for the consumer given by

$$R - e(p_0, V(p_1, R)).$$

The firm's profit loss is

$$\pi(p_0) - \pi(p_1 - t).$$

And the tax revenue collected by the government is (thus a welfare gain)

$$tx(p_1, R).$$

Thus the total welfare loss is

$$[R - e(p_0, V(p_1, R))] + [\pi(p_0) - \pi(p_1 - t)] - tx(p_1, R).$$

3 Optimal Commodities Taxation (Indirect Taxation)

3.1 Ramsey Model with a Representative Consumer

- J consumption goods, $j = 1, ..., J$;

- A representative consumer with utility function $u(\mathbf{x}, L)$ where $\mathbf{x} = (x_1, ..., x_J) \in R^J$ is her consumption vector and L is her labor supply; Each individual is endowed with 1 unit of labor.

- Productions of the J goods use labor, and are all linear:

$$x_j = \frac{L_j}{a_j}, j = 1, ..., J$$

  Without loss of generality, we can choose units of goods so that $a_j = 1$ for all j. Hence in equilibrium the *producer price* of the goods $p_j = w$ for all j where w is the equilibrium wage rate (normalized to 1 w.l.o.g).

- The government can set linear commodity tax $\mathbf{t} = (t_1, ..., t_J)$ and a linear wage income tax rate τ.

- The government needs to raise a total tax revenue of $R > 0$.

3.2 Ramsey Rule

First, note that we can without loss of generality set $\tau = 0$. To see this, suppose that $\tau \neq 0$. Note that the consumer's budget constraint will be

$$\sum_{j=1}^{J} (1 + t_j) x_j = (1 - \tau) L$$

where the left hand is the total expenditure on the consumption goods and the right hand side is the after tax wage income.

Now consider an alternative tax system in which

$$\begin{aligned} t'_j &= \frac{t_j + \tau}{1 - \tau}, j = 1, ..., J \\ \tau' &= 0. \end{aligned}$$

Clearly the individual's budget set is the same under $(t'_1, ..., t'_J, \tau')$ as that under $(t_1, ..., t_J, \tau)$. The government's tax revenue is also unchanged. To see this, note that the government's revenue is

$$\begin{aligned}
& \sum_{j=1}^{J} t_j x_j + \tau L \\
= & \sum_{j=1}^{J} t_j x_j + \tau \sum_{j=1}^{J} \frac{(1+t_j)}{1-\tau} x_j \\
= & \sum_{j=1}^{J} \left[t_j + \frac{\tau (1+t_j)}{1-\tau} \right] x_j \\
= & \sum_{j=1}^{J} t'_j x_j,
\end{aligned}$$

where the first equality exploits the consumer's budget constraint.

Ramsey Problem: The government sets $(t_1, ..., t_J)$, anticipating that the consumers will choose optimal consumption and labor supply plans in reaction to the government's tax system. The government's goal is to design a tax system that maximizes the representative consumer's utility subject to the revenue requirement of R.

To formally set up Ramsey's problem, we introduce some preliminary notations:

- Let $(\mathbf{x}^* (\mathbf{p},m), L^* (\mathbf{p},m)) \equiv (x_1^* (\mathbf{p},m), ..., x_J^* (\mathbf{p},m), L^* (\mathbf{p},m))$ be the solution to the consumer's problem when the tax vector is $\mathbf{t}$. That is,

$$\begin{aligned}
(\mathbf{x}^* (\mathbf{p},m), L^* (\mathbf{p},m)) &= \arg\max_{\{\mathbf{x},L\}} u(\mathbf{x},L) \\
\text{s.t. } \mathbf{p} \cdot \mathbf{x} - L &\leq m.
\end{aligned}$$

  $\mathbf{x}(\mathbf{p},m)$ is called the *Marshallian demand function.* We can interpret m as unearned income.

- The consumer's indirect utility function is

$$V(\mathbf{p},m) = u(\mathbf{x}^* (\mathbf{p},m), L^* (\mathbf{p},m)).$$

- The *compensated demand function* is the demand of the goods when utility is fixed at a certain level. Let $\left(\hat{\mathbf{x}}(\mathbf{p},u), \hat{L}(\mathbf{p},u)\right)$ be the consumer demand for the goods and labor supply when the price is $\mathbf{p}$, and the consumer's utility is fixed at u. That is,

$$\begin{aligned}
\left(\hat{\mathbf{x}}(\mathbf{p},u), \hat{L}(\mathbf{p},u)\right) &= \arg\min \{\mathbf{p} \cdot \mathbf{x} - L\} \\
\text{s.t. } u(\mathbf{x},L) &\geq u.
\end{aligned}$$

- With these notation, Ramsey's problem is

$$\begin{aligned} &\max_{\{\mathbf{t}\}} V(\mathbf{p},0) \\ \text{s.t.} \quad &\mathbf{t}\cdot\mathbf{x}^*(\mathbf{p},0) \geq R \end{aligned}$$

  where $\mathbf{p} = \mathbf{1}+\mathbf{t}$. The objective function is the representative consumer's indirect utility function and the constraint is the revenue requirement constraint. Note that in the constraint $\mathbf{x}^*(\mathbf{p})$ appears, that is, the consumer is best responding to the government tax system.

Derivation of Ramsey Rule:

If the tax system is $\mathbf{t} = (t_1, ..., t_J)$, then the *consumer price* for the goods will be $\mathbf{p} = (1+t_1, ..., 1+t_J)$.

Set the Lagrange of the maximization problem as

$$\begin{aligned} \mathcal{L} &= V(\mathbf{p},0) - \lambda\left[R - \mathbf{t}\cdot\mathbf{x}^*(\mathbf{p},0)\right] \\ &= V(\mathbf{p},0) - \lambda\left[R - \sum_{j=1}^{J} t_j x_j^*(\mathbf{p},0)\right] \end{aligned}$$

The FOC with respect to p_k is

$$\frac{\partial V(\mathbf{p},0)}{\partial p_k} + \lambda\left[x_k^*(\mathbf{p},0) + \sum_{j=1}^{J} t_j \frac{\partial x_j^*(\mathbf{p},0)}{\partial p_k}\right] = 0. \tag{1}$$

By Roy's identity,

$$\frac{\partial V(\mathbf{p},0)}{\partial p_k} = -\frac{\partial V(\mathbf{p},0)}{\partial m} x_k^*(\mathbf{p},0) \tag{2}$$

where $\partial V(\mathbf{p},0)/\partial m$ is the marginal utility of income. [Recall Roy's identity from your first year micro]

By Slutsky equation (which decomposes the effect of a price change on demand into income effect and substitution effect), we have

$$\frac{\partial x_j^*(\mathbf{p},0)}{\partial p_k} = \underbrace{\frac{\partial \hat{x}_j(\mathbf{p},u^*)}{\partial p_k}}_{\text{substition effect}} - \underbrace{x_k^*(\mathbf{p},0)\frac{\partial x_j^*(\mathbf{p},0)}{\partial m}}_{\text{income effect}}. \tag{3}$$

Plugging (2) and (3) into FOC (1), we obtain after some rearrangement that (neglecting the relevant arguments)

$$\sum_{j=1}^{J} t_j \left[\frac{\partial \hat{x}_j}{\partial p_k} - x_k^* \frac{\partial x_j^*}{\partial m}\right] = \left(\frac{\partial V/\partial m - \lambda}{\lambda}\right) x_k^*$$

Or,

$$\frac{\sum_{j=1}^{J} t_j \frac{\partial \hat{x}_j}{\partial p_k}}{x_k^*} = \frac{\partial V/\partial m - \lambda}{\lambda} + \sum_{j=1}^{J} \left(t_j \frac{\partial x_j^*}{\partial m} \right).$$

The final step, using the symmetry of the substitution effect, i.e.

$$\frac{\partial \hat{x}_j}{\partial p_k} = \frac{\partial \hat{x}_k}{\partial p_j},$$

we obtain the so-called **Ramsey Rule**:

$$\frac{\sum_{j=1}^{J} t_j \frac{\partial \hat{x}_k}{\partial p_j}}{x_k^*} = \frac{\partial V/\partial m - \lambda}{\lambda} + \sum_{j=1}^{J} \left(t_j \frac{\partial x_j^*}{\partial m} \right)$$

Interpretation of the Ramsey Rule: The left hand side is called "discouragement index for good k" To interpret this, note that

$$\hat{x}_k (1 + t, u^*) - \hat{x}_k (1, u^*) \approx \sum_{j=1}^{J} \frac{\partial \hat{x}_k}{\partial p_j} t_j$$

Thus,

$$\frac{\sum_{j=1}^{J} t_j \frac{\partial \hat{x}_k}{\partial p_j}}{x_k^*} \approx \frac{\hat{x}_k (1 + t, u^*) - \hat{x}_k (1, u^*)}{x_k^*}$$

is approximately equal to the proportional reduction in compensated demand for good k under the tax system $\mathbf{t}$.

The right hand side is *constant* across commoditities (does not depend on j).

Thus the Ramsey Rules states that the optimal tax system will have the property that the "discouragement index for all goods are equalized under the optimal tax system." Note that the Ramsey Rule does *not* call for uniform taxes across commodities.

Specical cases of the Ramsey Rule: Suppose that there is *no subsitution effect*, that is,

$$\frac{\partial \hat{x}_k}{\partial p_j} = 0 \text{ for all } j \neq k$$

and moreover, suppose that the *income effect is approximately zero*, then we obtain the so-called **Inverse Elasticity Rule:**

$$\frac{t_j}{p_j} = C \times \frac{1}{\epsilon_j^D}$$

where C is a constant given by

$$C = \frac{\partial V/\partial m - \lambda}{\lambda}$$

and ϵ_j^D is good-$j's$ own price elasticity of demand.

3.3 Extenstion to the Heterogenous Consumer Case

- Suppose that there are H consumers $h = 1, ..., H$;
- Consumer $h's$ utility function is $u^h\left(\mathbf{x}^h, L^h\right)$;
- Given price vector $\mathbf{p}$, and unearned income m^h, let

$$\left(\mathbf{x}^{*h}\left(\mathbf{p}, m^h\right), L^{*h}\left(\mathbf{p}, m^h\right)\right)$$

denote consumer $h's$ demand for the consumption goods and labor supply; and let $V^h\left(\mathbf{p}, m^h\right)$ denote her indirect utility.

- Imagine a Bergson-Samuelson social welfare function

$$\chi(\mathbf{p}, \mathbf{m}) = W\left(V^1\left(\mathbf{p}, m^1\right), ..., V^H\left(\mathbf{p}, m^H\right)\right)$$

where $\mathbf{m} = \left(m^1, m^2, ..., m^H\right)$ are the vector of unearned incomes for the H consumers.

- The government's problem is

$$\begin{aligned} & \max \chi(\mathbf{p}, \mathbf{m}) \\ \text{s.t.} \quad & \sum_{h=1}^{H} \sum_{j=1}^{J} \underbrace{(p_j - 1)}_{t_j} x_j^{*h}\left(\mathbf{p}, m^h\right) \geq R \end{aligned}$$

- The FOC with respect to p_k is given by

$$\sum_{h=1}^{H} \frac{\partial \chi}{\partial V^h} \frac{\partial V^h}{\partial p_k} = -\lambda \sum_{h=1}^{H} \left[x_k^{*h}\left(\mathbf{p}, m^h\right) + \sum_{j=1}^{J} t_j \frac{\partial x_j^{*h}\left(\mathbf{p}, m^h\right)}{\partial p_k} \right]$$

where λ is the Lagrange multiplier of the government budget constraint.

By Roy's identigy,

$$\begin{aligned} \frac{\partial V^h}{\partial p_k} &= -\frac{\partial V^h}{\partial m} x_k^{*h} \\ &= -\alpha^h x_k^{*h} \end{aligned}$$

where $\alpha^h \equiv \partial V^h / \partial m$ is consumer $h's$ marginal utility of income.

From Slutsky equation, we have

$$\frac{\partial x_j^{*h}\left(\mathbf{p}, m^h\right)}{\partial p_k} = \underbrace{\frac{\partial \hat{x}_j^h}{\partial p_k}}_{\text{substition effect}} \underbrace{- x_k^{*h}\left(\mathbf{p}, m^h\right) \frac{\partial x_j^{*h}\left(\mathbf{p}, m^h\right)}{\partial m}}_{\text{income effect}}$$

With these two formulas, we obtain

$$\sum_{h=1}^{H} \frac{\partial \chi}{\partial V^h} \alpha^h x_k^{*h} = \lambda \sum_{h=1}^{H} \left\{ x_k^{*h}\left(\mathbf{p}, m^h\right) + \sum_{j=1}^{J} t_j \left[\frac{\partial \hat{x}_j^h}{\partial p_k} - x_k^{*h}\left(\mathbf{p}, m^h\right) \frac{\partial x_j^{*h}\left(\mathbf{p}, m^h\right)}{\partial m} \right] \right\}$$

equivalently

$$\sum_{j=1}^{J} t_j \left(\sum_{h=1}^{H} \frac{\partial \hat{x}_j^h}{\partial p_k} \right) = \frac{\sum_{h=1}^{H} \frac{\partial \chi}{\partial V^h} \alpha^h x_k^{*h}}{\lambda} - \sum_{h=1}^{H} x_k^{*h}\left(\mathbf{p}, m^h\right) + \sum_{h=1}^{H} x_k^{*h}\left(\mathbf{p}, m^h\right) \sum_{j=1}^{J} t_j \frac{\partial x_j^{*h}\left(\mathbf{p}, m^h\right)}{\partial m}$$

Finally, use the symmetry of the substitution matrix, i.e.,

$$\frac{\partial \hat{x}_j^h}{\partial p_k} = \frac{\partial \hat{x}_k^h}{\partial p_j}$$

we obtain:

$$\sum_{j=1}^{J} t_j \left(\sum_{h=1}^{H} \frac{\partial \hat{x}_k^h}{\partial p_j} \right) = \frac{\sum_{h=1}^{H} \frac{\partial \chi}{\partial V^h} \alpha^h x_k^{*h}}{\lambda} - \sum_{h=1}^{H} x_k^{*h}\left(\mathbf{p}, m^h\right) + \sum_{h=1}^{H} x_k^{*h}\left(\mathbf{p}, m^h\right) \sum_{j=1}^{J} t_j \frac{\partial x_j^{*h}\left(\mathbf{p}, m^h\right)}{\partial m}$$

A much better way to write the above formula is as follows:

- Write $X_k = \sum_{h=1}^{H} x_k^{*h}\left(\mathbf{p}, m^h\right)$ as the total demand for good k.
- Adopt the following shortcut:

$$\beta^h = \frac{\frac{\partial \chi}{\partial V^h} \alpha^h}{\lambda} + \sum_{j=1}^{J} t_j \frac{\partial x_j^{*h}\left(\mathbf{p}, m^h\right)}{\partial m}$$

  The first term of β^h is the social marginal utility of income of consumer h, divided by λ - the cost of budget resources for the government. The second term is the increase in tax revenue collected on h when his income is increased by one unit. β^h measures what is called "the net social marginal utility of income for consumer h."

Using the above two pieces of short-cut notation, we can write the formula as

$$\sum_{j=1}^{J} t_j \left(\sum_{h=1}^{H} \frac{\partial \hat{x}_k^h}{\partial p_j} \right) = -X_k \left(1 - \sum_{h=1}^{H} \beta^h \frac{x_k^{*h}}{X_k} \right)$$

Let $\bar{\beta} = \sum_{h=1}^{H} \beta^h / H$ denote the average of β^hs over the consumers; and $\bar{X}_k = X_k / H$ as the average consumption of good k.

Let θ_k the empirical covariance between β^h and $h's$ consumption of good k, i.e.,

$$\theta_k = \text{cov}\left(\frac{\beta^h}{\bar{\beta}}, \frac{x_k^{*h}}{\bar{X}_k}\right) = \frac{1}{H}\sum_{h=1}^{H}\left(\frac{\beta^h}{\bar{\beta}} - 1\right)\left(\frac{x_k^{*h}}{\bar{X}_k} - 1\right).$$

- A positive θ_k implies that good k are heavily consumed by agents with a higher β^h, i.e. by agents with a high net social marginal utility of income.

- One can show that

$$\theta_k = \sum_{h=1}^{H}\frac{\beta^h x_k^{*h}}{\bar{\beta} X_k} - 1$$

Thus the formula can be written as

$$-\frac{\sum_{j=1}^{J} t_j\left(\sum_{h=1}^{H}\frac{\partial \hat{x}_k^h}{\partial p_j}\right)}{X_k} = 1 - \bar{\beta} - \bar{\beta}\theta_k.$$

This is called the **Generalized Ramsey Rule.**

The left hand size is called the "*discouragement index for good k under the tax system* **t**." The numerator, as before, measure the total reduction in the demand (over all consumers) on good k as a result of the tax system; and the denominator is the total demand for good k. Hence the LHS is the percentage reduction in the total demand for good as a result of the tax system.

The RHS is called "*the redistributive factor of good k*." Note also that in a one consumer economy, θ_k is obviously zero, therefore we can get back the previous Ramsey rule. Note that it is **not** a constant in this heterogenous consumer economy.

Implications of the Generalized Ramsey Rule:

The above formula suggests that in a heterogenous-agent economy, the optimal commodity tax system should discourage less the consumption of these goods that have a large and positive θ_k. Recall a positive θ_k implies that good k are heavily consumed by agents with a higher β^h, i.e. by agents with a high net social marginal utility of income. But who are those agents with higher β^h : from the definition of β^h, these are the agents with a higher $\partial\chi/\partial V^h$ (i.e. those that are previleged by the government in its objective function); or with higher $\partial x_k^{*h}/\partial m$ (i.e. agents with higher income effect in their demands, typically poor people). Therefore the generalized Ramsey rule implies that under the optimal commodity tax system, the consumption of the goods that the poor buy more should be less discouraged.

4 Optimal Income Taxation (Direct Taxation)

- Classic Paper: Mirrlees (1971): "An Exploration in the Theory of Optimal Income Taxation." Review of Economic Studies, Vol. 38, 175-208.

The goal of this literature is to explore some of the consierations that enter the determination of the extent of redistributive income taxation. In particular, how is the optimal structure of income taxation influenced by differences in distributional objectives, by the responsiveness of labor supply to taxation, and by the magnitude of differences in the pre-tax incomes.

4.1 A Simple Benchmark with No Disincentive Effects

- There is a continuum of agents. Agents differ in their earning ability w distributed according to a CDF $F(w)$;
- Agent of type w has a utility function $U_w(\cdot)$;
- The pre-tax income for an agent with earning ability w is denoted by $Z(w)$;
- Write the tax levied on type-w agent as $T(w)$;
- The government's tax revenue constraint is R;
- If the government wants to maximize a utilitarian social welfare

$$\max_{T(\cdot)} \int_0^\infty U_w(Z(w) - T(w))\, dF(w)$$
$$\text{s.t.} \quad \int_0^\infty T(w)\, dF(w) = R.$$

- If $U_w' > 0$ and $U_w'' < 0$, then the solution will have

$$U_w'(Z(w) - T(w)) = \text{constant for all} w.$$

  That is, the taxes will be chosen so that the marginal utility of the consumption at the after-tax income level is equalized across individuals.

- Under the special case that U_w is the same for all w, then it requires equalization of the after-tax incomes under the optimal tax structures.

Problems:

1. It does not take into account of disincentive effects of taxes (labor supply is not part of the model);

2. Utilitarian social welfare framework may be inadequate. First, there is the problem with interpersonal comparison of utilities; second, even if one accepts interpersonal comparison of utilities, like Rawles, the appropriate maximand for the government may not be the sum of individual utilities. Rawls, for example, advocate a "maxi-min" criterion, i.e., the guiding principle for the government should be the maximization of the welfare of the least fortunate person.

3. Tax instruments: Suppose that the pre-tax income is generated by working $L(w)$ hours, i.e. $Z(w) = w.L(w)$. Under the assumption that $L(w)$ is not affected by tax (i.e. no disincentive effect), the equalization of after-tax incomes can be achieved by lump sum tax $(Z - \tau(w))$, a wage tax $[w - t(w)]L$, or by an income tax $Z - T(Z)$.

4.2 Optimal Linear Income Tax with Disincentive Effect

We first analyze the optimal linear income tax with endogenous labor supply. This is a useful case to analyze because the structure of the optimal linear income tax is similar to that of the optimal commodity taxes.

Good reference to this:

- Sheshinski, Eytan. "The Optimal Linear Income Tax." Review of Economic Studies, 1972, 297-302.

- Dixit, Avinash and Agnar Sandmo. "Some Simplified Formulae for Optimal Income Taxation." Scandinavian Journal of Economics.

The Model

- A continuum of agents with earning ability $w \sim F(w)$ with a support $[\underline{w}, \infty]$. If a type-w agent works $L(w)$ hours, her income will be $Z(w) = wL(w)$. $L(w)$ is endogenous.

- Suppose that all agents have the **same** utility function $u(C, L)$ where C is the numeraire good, and L is the labor supply. u has the usual desirable properties.

- The linear income tax consists of two elements, a proportional wage income tax at a rate of t; and a head-tax (or transfer) G.[2] The government needs to raise a total tax revenue R.

- Given (t, G), an agent with earning ability w solves

$$\max_{\{L \geq 0\}} u\left((1-t) wL + G, L\right).$$

The FOC is either

$$(1-t) w u_1' + u_2' = 0$$

or

$$(1-t) w u_1'(G, 0) + u_2'(G, 0) \leq 0 \text{ and } L = 0.$$

It is easy to see that there exists a threhold w_0 such that

$$L^*(w) = 0 \text{ for } w \leq w_0; \text{ and } L^*(w) > 0 \text{ for } w > w_0.$$

Of course this threshold w_0 depends on the income tax rate t and transfer G. [For notational simplicity below, we supress this dependence in write $L^*(w; t, G)$.]

Let $V((1-t) w, G)$ [shortcut $V(w)$] as the indirect utility function of the type-w agent under a tax policy (t, G).

- The resource constraint for the economy is

$$\int_{\underline{w}}^{\infty} C^*(w)\, dF(w) + R = \int_{\underline{w}}^{\infty} wL^*(w)\, dF(w)$$

Noting that $C^*(w) = (1-t) wL^*(w) + G$, we have

$$G + R = t \int_{\underline{w}}^{\infty} wL^*(w)\, dF(w).$$

- The government's problem is to maximize a Bergson-Samuelson social welfare function

$$\max_{\{t,G\}} \int_{\underline{w}}^{\infty} \chi(V(w))\, dF(w)$$
$$\text{s.t. } G + R = t \int_{\underline{w}}^{\infty} wL^*(w)\, dF(w).$$

[2] Aside: the difference between a lump-sum tax and a head-tax. A lump-sum tax is any tax that is independent of what the agent does (hence it does not create incentive effect except possibly through income effect). A lump-sum tax can still be individual specific (for example, the elderly receives a larger lump-sum check than the working-age). A head tax is a special case of a lump-sum tax that has to be the same for everyone, hence the name "head" tax.

Analysis:

Form the Lagrangian,

$$\mathcal{L} = \int_{\underline{w}}^{\infty} \{\chi (V (w)) + \lambda [twL^* (w) - G - R]\} dF (w)$$

The first order conditions with respect to G and t are respectively

$$\begin{aligned} \int_{\underline{w}}^{\infty} \left\{\chi' \frac{\partial V}{\partial G} + \lambda \left[tw \frac{\partial L^*}{\partial G} - 1\right]\right\} dF (w) &= 0 \\ \int_{\underline{w}}^{\infty} \left\{\chi' \frac{\partial V}{\partial t} + \lambda \left[wL^* (w) + tw \frac{\partial L^*}{\partial t}\right]\right\} dF (w) &= 0 \end{aligned}$$

[Note that these are simply necessary first order conditions, and there is no guarantee yet that solutions to the first order conditions is optimal. But examining the implications of the necessary FOC is useful to know the properties of the optimal solution.]

Using Slutsky equation, we can write

$$\frac{\partial L^*}{\partial t} = -w \frac{\partial \hat{L}}{\partial w} - wL^* \frac{\partial L^*}{\partial m}$$

where $\partial \hat{L} / \partial w$ is the substitution term (compensated response of labor to the marginal net wage), it is multiplied by $-w$ because a marginal increase of tax by Δ decreases net wage by a Δw; the term $-wL^* \partial L^* / \partial m$ is the income effect.

By Roy's identity,

$$\begin{aligned} \frac{\partial V}{\partial t} &= -\frac{\partial V}{\partial m} wL^* \\ &= -\alpha w L^* \end{aligned}$$

where $\alpha = \partial V / \partial m$ is the private marginal utility of income for a type-w agent. Also,

$$\frac{\partial L^*}{\partial G} = \frac{\partial L^*}{\partial m}, \text{ and } \frac{\partial V}{\partial G} = \frac{\partial V}{\partial m} = \alpha$$

because G is just a head transfer.

With these, we can rewrite the FOCs as

$$\int_{\underline{w}}^{\infty} \left[\chi' \frac{\alpha}{\lambda} + tw \frac{\partial L^*}{\partial m} - 1\right] dF (w) = 0 \tag{4}$$

$$\int_{\underline{w}}^{\infty} wL^* \left\{\frac{\chi' \alpha}{\lambda} + tw \frac{\partial L^*}{\partial m} - 1 + \frac{tw}{L^*} \frac{\partial \hat{L}}{\partial w}\right\} dF (w) = 0 \tag{5}$$

Now, as in the optimal commodity taxation case, define

$$\beta(w) = \frac{\chi' \alpha}{\lambda} + tw \frac{\partial L^*}{\partial m}$$

as the net social marginal valuation of income for a type-w agent. Thus the two necessary conditions are reduced to

$$\bar{\beta} = 1 \quad (6)$$

$$t = \frac{-\mathrm{cov}(\beta(w), Z(w))}{\int_{\underline{w}}^{\infty} Z(w) \epsilon_L^S(w) \, dF(w)} \quad (7)$$

where $\mathrm{cov}(\beta(w), Z(w))$ is the covariance between $\beta(w)$ and $Z(w)$, and

$$\epsilon_L^S(w) = \frac{w}{\hat{L}} \frac{\partial \hat{L}}{\partial w} = \frac{w}{L^*} \frac{\partial \hat{L}}{\partial w}$$

is the compensated labor supply elasticity. While the first condition $\bar{\beta} = 1$ is obvious, the second can be shown as follows: Condition (5) can be re-arranged as

$$\int_{\underline{w}}^{\infty} \underbrace{Z(w)}_{=wL^*(w)} \left[\beta(w) - \bar{\beta} + t\epsilon_L^S(w)\right] dF(w) = 0$$

Because

$$\mathrm{cov}(\beta(w), Z(w)) = \int_{\underline{w}}^{\infty} Z(w) \left[\beta(w) - \bar{\beta}\right] dF(w)$$

we have

$$\mathrm{cov}(\beta(w), Z(w)) = -t \int_{\underline{w}}^{\infty} Z(w) \epsilon_L^S(w) \, dF(w).$$

Interpretations:

- Condition (6) says that the head-transfer element of the tax system should be adjusted such that $\bar{\beta}$, the net social marginal valuations of the transfer of a dollar of income, should on avearge be equal to the cost of one dollar;

- Condition (7) neatly captures the tradeoff of efficiency and equity in designing the optimal income tax. The numerator caputures equity, while the denominator captures efficiency. To see this, note that $-\mathrm{cov}(\beta(w), Z(w))$ will be positive if $\beta(w)$, the social marginal valuation of income, is higher for agents with lower pre-tax earnings. This clearly captures the equity. The denominator is the labor supply response to the income tax. The optimal income tax rate would adjust to keep this balance between equity and efficiency.

4.3 An Example

- Consider Cobb-Douglas utility function

$$u(C, L) = \gamma \ln C + (1-\gamma) \ln (1-L)$$

  where $\gamma \in (0,1)$;

- It is easy to see that given tax system (t, G), the optimal labor supply $L^*(w)$ will be

$$\begin{aligned} L^*(w) &= \gamma - \frac{(1-\gamma) G}{w(1-t)} \text{ if } w \geq \frac{1-\gamma}{\gamma} \frac{G}{1-t} \equiv w_0 \\ L^*(w) &= 0 \text{ otherwise.} \end{aligned}$$

  Thus the net social marginal valuation of income $\beta(w)$ is given by

$$\begin{aligned} \beta(w) &= \chi' \frac{\alpha}{\lambda} - \frac{t}{1-t} (1-\gamma) \text{ for } w \geq w_0 \\ &= \chi' \frac{\alpha}{\lambda} \text{ otherwise.} \end{aligned}$$

- The exact form of optimal tax system depend on the function form χ; and the distribution F.

4.4 General Analysis of the Optimal Income Taxes