Auerbach Hines 2002 Taxation and Economic Efficiency

Xu Jiang

Outline

Overview

Measures of tax distortion

Design of optimal tax - the Ramsey problem

Optimal design of individual income taxes

Optimal design of capital income tax

The purpose of optimal taxation

- Non lump-sum taxes necessarily generates distortions as it affects real decisions
- But essential because of externalities or the provision of public goods or redistribution of income (to maximize social welfare)
- Holy grail of optimal taxation: the least distortionary way (or the least excess burden) of designing optimal tax systems that can achieve those objectives
- We first need a measure of the distortion, or the excess burden, induced by tax.

Outline

Defining tax distortion based on (change in) Marshallian measure of the consumer surplus

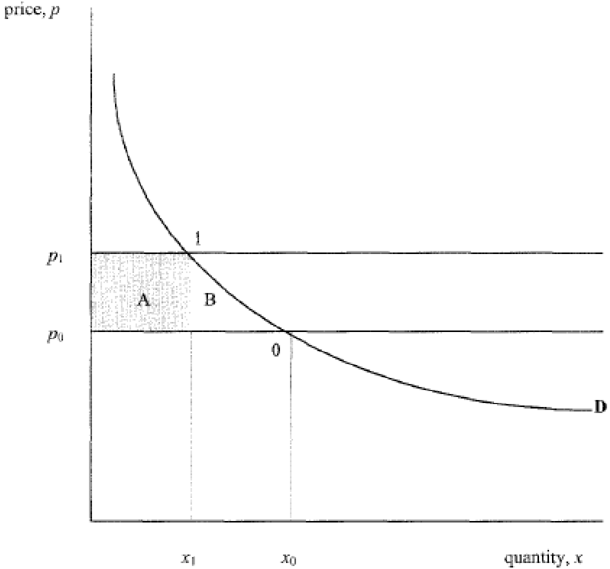
- Consider a simple good: without taxes the market-clearing prices and quantities are p_0 and x_0.
- Imposing a sales tax increases prices to p_1 and reduces demand to x_1.

- The distortion is the “Harberger” triangle.

Marshallian consumer surplus based tax distortion - example

- Consider an inverse demand function $p(x) = 10 - x$.
- $p_0 = 5$ results in $x_0 = 10 - 5 = 5$ and the consumer surplus is $5 \times 5 = 25$.
- Now consider imposing a sales tax of 30%, resulting in $p_1 = 5 \times (1 + 30\%) = 8$ and $x_1 = 10 - 8 = 2$.
- The new consumer surplus is $8 \times 2 = 16$ and the tax revenue collected by the government is $16 \times 30\% = 4.8$
- The distortion is therefore $25 - (16 + 4.8) = 4.2$.
- Intuition: the reduced demand for the taxed goods results in a smaller tax base and thus a reduction in total welfare.

Marshallian consumersurplus based tax distortion - Graphical Illustration (Figure 2.1)



Issues with Marshallian consumer surplus based tax distortion

- Marshallian consumer surplus not uniquely defined when there is more than one commodity and more than one price change (Auerbach 1985), that is, it is path dependent. Reason: the definition of Marshallian consumer surplus does not come directly from underlying consumer preferences.
- Therefore need alternative measures that are not path dependent.

More on path dependence

- For a single good, consumer surplus is defined as $\Delta S = -\int_{p_0}^{p_1} x(p)dp$ where $x(p)$ is the demand function.
- Suppose there are two goods and both prices change. If we change the price in market 1 first, the change in surplus is

$$\Delta S_1 = -\int_{p_0^1}^{p_1^1} x^1(p^1, p_0^2)dp^1 - \int_{p_0^2}^{p_1^2} x^2(p_1^1, p^2)dp^2,$$

  whereas if we change the price in market 2 first, the change in surplus is

$$\Delta S_2 = -\int_{p_0^1}^{p_1^1} x^1(p^1, p_1^2)dp^1 - \int_{p_0^2}^{p_1^2} x^2(p_0^1, p^2)dp^2.$$

- For small changes, that is, $p_1^2 = p_0^2 + dp_0^2$ and $p_1^1 = p_0^1 + dp^1$, we can calculate $\Delta S_2 - \Delta S_1$ to be

$$\Delta S_2 - \Delta S_1 = [\frac{\partial x^2(p_0^1, p_0^2)}{\partial p^1} - \frac{\partial x^1(p_0^1, p_0^2)}{\partial p^2}]dp^1 dp^2,$$

  which is zero if and only if $\frac{\partial x^2(p_0^1,p_0^2)}{\partial p^1} - \frac{\partial x^1(p_0^1,p_0^2)}{\partial p^2} = 0$.

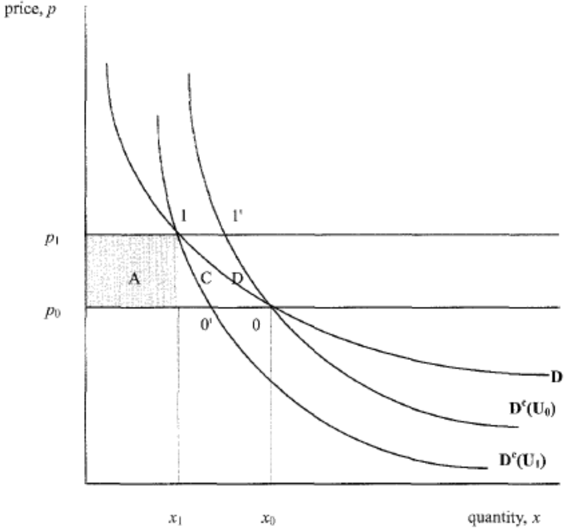
Hicksian consumer-surplus based tax distortion

- This measure resolves the issue, as Hicksian measure holds utility rather than income constant (Slutsky matrix is symmetric, i.e., $\frac{\partial x^2(p_0^1,p_0^2)}{\partial p^1}\Big|_{\text{utility constant}} - \frac{\partial x^1(p_0^1,p_0^2)}{\partial p^2}\Big|_{\text{utility constant}} = 0$).
- When price increases from p_0 to p_1, consider the hypothetical example where the government compensates consumers so that they remain at original utility levels. Calculate tax revenue collected based on this hypothetical case.
- Tax distortion is defined as the difference between the amount of compensation to keep the consumers' utilities constant and the tax revenue collected.

Hicksian consumer surplus- based tax distortion - example

- Consider an inverse demand function $p(x) = 10 - x$, with the consumer's utility function being $u(x) = x$.
- Without any tax, $p_0 = 5$ results in $x_0 = 10 - 5 = 5$ and the consumer's utility is $u(x_0) = 5$.
- Now consider imposing a sales tax of 30%, resulting in $p_1 = 5 \times (1 + 30\%) = 8$ and $x_1 = 10 - 8 = 2$. The current utility is $u(x_1) = 2$.
- For the consumer to achieve the same utility level as that before tax, the government needs to compensate the consumer $(5 - 2) \times 8 = 24$ so the consumer can still consumer 5 units.
- Based on the consumer still consuming 5 units, the tax collected is $5 \times 8 \times 30\% = 12$.
- The distortion is therefore $24 - 12 = 12$.
- Intuition: to get the consumer back to the same utility level, the government needs to fully compensate for the lost demand but only collect a fraction as sales tax revenue.

Hicksian consumer-surplus based tax distortion - Graphical Illustration (Figure 2.2)



Outline

What is the Ramsey problem

- Government needs to raise an exogenous and fixed amount of tax revenue.
- Government imposes the tax on commodities, like a sales tax.
- Assume a population of identical individuals, which can be reduced to the analysis of a representative individual.
- The government designs the optimal tax system to minimize the distortion, subject to raising the fixed revenue.

Ramsey problem setup - three commodities example

- The representative consumer maximizes utility $U(x_0, x_1, x_2)$ subject to a budget constraint $\sum_{i=0}^{2} p_i x_i \leq y$ where y is some exogenous endowment income. The maximization problem generates an indirect utility function $V(p_0, p_1, y) = \max_{\sum_{i=0}^{2} p_i x_i \leq y} U(x_0, x_1)$.
- The government imposes taxes on the commodities, t_0, t_1 and t_2, resulting in $p_i = q_i + t_i$ for $i = 0, 1, 2$, to collect a revenue of R, where $q_i s$ are producer prices. Choosing t_i is thus equivalent to choosing p_i given fixed q_i.
- The government's optimization problem is thus

$$\max_{p_0, p_1, p_2} V(p_0, p_1, p_2, y) \text{ s.t. } \sum_{i=0}^{2} (p_i - q_i) x_i \geq R. \tag{1}$$

Solving the Ramsey problem

- Problem 1 is equivalent to

$$\min_{p_0,p_1,p_2} y - E[q_0, q_1, q_2, V(p_0, p_1, p_2, y)] - R \text{ s.t. } \sum_{i=0}^{2}(p_i - q_i)x_i \geq R.$$

  where $E[q_0, q_1, q_2, V(p_0, p_1, p_2, y)]$ is the expenditure function (i.e., $p_0x_0 + p_1x_1 + p_2x_2$) when the consumer's utility level is $V(p_0, p_1, p_2, y)$.
- Since $y \equiv E[p_0, p_1, p_2, V(p_0, p_1, p_2, y)]$, $y - E[q_0, q_1, q_2, V(p_0, p_1, p_2, y)] - R$ is the Hicksian consumer surplus-based tax distortion.
- Clearly any solution that results in zero distortion is optimal.

Solving the Ramsey problem, continued

- Denote $\overrightarrow{p} = (p_0, p_1, p_2)$ and $\overrightarrow{q} = (q_0, q_1, q_2)$.
- $p_i = \phi q_i$ (i.e., uniform sales taxes) such that $(\phi - 1)(\sum_{i=0}^{2} q_i x_i) = R$ is thus optimal as

$$\begin{aligned}
& y - E[\overrightarrow{q}, V(\overrightarrow{q}, y)] - R \\
= \;& E[\overrightarrow{p}, V(\overrightarrow{p}, y)] - E[\overrightarrow{q}, V(\overrightarrow{q}, y)] - (\phi - 1)(\sum_{i=0}^{2} q_i x_i) \\
= \;& E[\phi\overrightarrow{q}, V(\overrightarrow{p}, y)] - E[\overrightarrow{q}, V(\overrightarrow{q}, y)] - (\phi - 1)(\sum_{i=0}^{2} q_i x_i) \\
= \;& \phi E[\overrightarrow{q}, V(\overrightarrow{q}, y)] - E[\overrightarrow{q}, V(\overrightarrow{q}, y)] - (\phi - 1)(\sum_{i=0}^{2} q_i x_i) \\
= \;& (\phi - 1)(\sum_{i=0}^{2} q_i x_i) - (\phi - 1)(\sum_{i=0}^{2} q_i x_i) = 0.
\end{aligned}$$

Solving the Ramsey problem, continued

- Suppose such proportional tax is not allowed. Then WLOG normalize the taxes on one of the commodities to zero and WLOG assume this is also the numeraire good, that is, $q_0 = p_0 = 1$.
- We can now rewrite problem 1 as

$$\max_{p_1, p_2} V(p_1, p_2, y) \text{ s.t. } (p_1 - q_1)x_1 + (p_2 - q_2) \geq R$$

- Denote μ as the Lagrangian multiplier associated with the constraint. FOC results in

$$-\frac{\partial V}{\partial y} x_1 + \mu(x_1 + t_1 \frac{dx_1}{dp_1} + t_2 \frac{dx_2}{dp_1}) = 0,$$

and

$$-\frac{\partial V}{\partial y} x_1 + \mu(x_2 + t_1 \frac{dx_1}{dp_2} + t_2 \frac{dx_2}{dp_2}) = 0.$$

Solving the Ramsey problem, continued

- Using the Slutsky equation

$$\frac{dx_i}{dp_j} = \frac{dx_i}{dp_j}\Big|_{y \text{ constant}} - \frac{dx_i}{dy}x_j,$$

the first order conditions are equivalent to

$$t_1\frac{dx_1}{dp_1} + t_2\frac{dx_2}{dp_1} = -\frac{\mu - \alpha}{\mu}x_1 \tag{2}$$

and

$$t_1\frac{dx_1}{dp_2} + t_2\frac{dx_2}{dp_2} = -\frac{\mu - \alpha}{\mu}x_2 \tag{3}$$

where $\alpha \equiv \frac{\partial V}{\partial y} + \mu\left(t_1\frac{dx_1}{dy} + t_2\frac{dx_2}{dy}\right)$ is the "social" marginal utility of income. The excess burden is captured by how much μ is larger over α.

Properties of optimal taxes

- Equation (2) divided by equation (3) results in

$$\frac{t_1}{t_2} = \frac{-\frac{dx_2}{dp_2}x_1 + \frac{dx_1}{dp_2}x_2}{-\frac{dx_1}{dp_1}x_2 + \frac{dx_2}{dp_1}x_1}.$$

- Denote $\theta_i = \frac{t_i}{p_i}$ as the tax rate on good i and using the fact that $\sum_i p_i \frac{dx_i}{dp_j} = \frac{dy}{dp_j} = 0$ results in

$$\frac{\theta_1}{\theta_2} = \frac{\varepsilon_{20} + \varepsilon_{21} + \varepsilon_{12}}{\varepsilon_{10} + \varepsilon_{21} + \varepsilon_{12}},$$

where ε_{ij} is the compensated cross-price elasticity of demand for good i with respect to the price of good j, which is equal to $\frac{\frac{dx_i}{x_i}}{\frac{dp_j}{p_j}} = \frac{dx_i}{dp_i}\frac{p_j}{x_i}$.

Properties of optimal taxes, continued

- Proportional tax is optimal if and only if $\varepsilon_{20} = \varepsilon_{10}$.
- If $\frac{dx_1}{dp_2} = \frac{dx_2}{dp_1} = 0$ (which may not be unreasonable if thinking of good 0 as labor, goods 1 and 2 as consumption goods), then $\theta_i \sim \frac{1}{\varepsilon_{i0}}$, that is, higher tax for goods that have lower cross-price elasticity to e.g., labor.

Outline

Overview

Measures of tax distortion

Design of optimal tax - the Ramsey problem

Optimal design of individual income taxes

Optimal design of capital income tax

Simplest case

- Assume one consumption good, only income coming from wages, individuals have common utility functions $U(c, l)$ with respect to consumption and leisure.
- Individuals have different abilities, as modelled in a reduced form - their different wages w^i.
- The government needs to raise R by charging T_i to individual $i \in \{1, 2, .., H\}$, that is, $c^i = w^i L^i - T^i$.
- The government's problem is

$$\max_{\{T_i\}_{i=1}^{H}} W(V^1(w^1, -T^1), ..., V^H(w^H, -T^H)) \text{ s.t. } \sum_h T^h \geq R.$$

- Denote μ as the Lagrangian multiplier of the constraint, then FOC results in $\frac{\partial W}{\partial V^i} \frac{\partial V^i}{\partial T^i} = \mu \ \forall i$, i.e., the marginal social utility of income is the same across all individuals.

Simplest case, continued

- Suppose W is a sum of individual's utility functions, then $\frac{\partial W}{\partial V^i} = 1$, implying that $\frac{\partial V^i}{\partial T^i}$ is constant across individuals, which in turn implies that marginal utility of consumption is constant across individuals.
- The marginal utility of leisure, however, is proportional to w^h.
- This implies that all individuals have the same level of consumption, and individuals with the lowest ability have the lowest wage, highest amount of leisure and therefore the highest utility.
- Clearly not incentive compatible when each individual's ability is not observable, as all other individuals want to pretend they have the lowest ability.

Optimal income taxation when individuals' abilities are unobservable (Mirrlees 1971 RES)

- For simplicity assume that government revenue target $R = 0$ (i.e., the government is only interested in redistribution).
- Assume individual skill/wage unobservable, denoted by w, but the government is aware of the distribution of w, $f(w)$.
- The government's problem (a mechanism design problem):

$$\max_{c(w),y(w)} \int_w G(U(w))f(w)dw$$

$$\text{s.t.} \int_w (c(w) - y(w))f(w)dw \leq 0$$

and $(c(w), y(w))$ being incentive compatible.

where $c(w)$ and $y(w)$ are the levels of consumption and income if the individual claims to have ability w (i.e., income tax would be $y(w) - c(w)$) and $U(w)$ is the individual's utility function $U(c(w), 1 - l) = U(c(w), 1 - \frac{y(w)}{w})$.

What it means to be incentive compatible

- Incentive compatibility means that any individual who knows his ability w has no incentive to choose $(c(w'), y(w'))$ over $(c(w), y(w))$, that is,

$$U(c(w), 1 - \frac{y(w)}{w}) \geq U(c(w'), 1 - \frac{y(w')}{w})\ \forall w, w'.$$

- Let w' be sufficiently close to w, then FOC results in

$$\frac{\partial U}{\partial c}\frac{dc}{dw} + \frac{\partial U}{\partial y}\frac{dy}{dw} = 0.$$

- This results in

$$\frac{dU}{dw} = \frac{\partial U}{\partial w} = U_2 \frac{y}{w^2} = U_2 \frac{L}{w}$$

Solving for the optimization problem

- Choosing $(c(w), y(w))$ is equivalent to choosing $(u(w), L(w))$ as $u = U(c, 1-L)$ and $y = wL$.
- Denote μ as the Lagrangian multiplier of the feasibility constraint and η as the incentive compatibility constraint, the optimization problem can be expressed as an optimal control problem with the Hamiltonian:

$$H = [G(u) - \mu(c(L,u) - y(L,u))]f(w) - \eta(w)U_2(L,u)\frac{L}{w}.$$

Solving for the optimization problem, continued

$$\frac{\partial H}{\partial L} = 0,$$

which is equivalent to

$$-\mu[\frac{\partial c}{\partial L}|_u - \frac{\partial y}{\partial L}|_u]f(w) - \eta(w)[\frac{\partial U_2}{\partial L}|_u \frac{L}{w} + \frac{U_2}{w}] = 0,$$

which is equivalent to

$$\frac{t}{1-t} = (\frac{U_1 \eta}{\mu})\frac{\Psi}{wf(w)}, \tag{4}$$

where $\Psi \equiv \frac{\partial U_2}{\partial L}|_u \frac{L}{U_2} + 1$, $t = \frac{\partial(y(w)-c(w))}{\partial y(w)} = 1 - \frac{\partial c(w)}{\partial y(w)}$ is the marginal tax rate, and we use $\frac{U_2}{U_1} = w(1-t)$ (individual maximization of $U(c(w), 1 - \frac{y(w)}{w})$ and taking derivative w.r.t. to $y(w)$).

Optimal marginal tax rate on labor income

- Equation (4) states that the optimal marginal tax rate on labor income decreases in $wf(w)$: the more effective labor supply is subject to the marginal tax rate at w, the less the rate should be to minimize distortion.
- Optimal marginal tax rate increases in Ψ, which can be shown to be decreasing in labor supply elasticity given $U(c, l) = c - v(1 - l) = c - v(L)$, which is $\frac{\partial \ln v'(L)}{\partial L} = \frac{v''(L)L}{v'(L)}$. Therefore, higher labor supply elasticity results in a lower marginal tax rate, as higher labor supply elasticity results in more tax distortion.
- In general very little can be said about the general shape of the optimal marginal tax rate as it depends on the social welfare function and the wage distribution.

Outline

Optimal capital income tax should be zero in the steady state of an intertemporal model

- First reported by Chamley (1986, Econometrica) and Judd (1985, Journal of Public Economics)
- One of the classical results in modern optimal tax theory
- Surprising in light of the standard Ramsey intuition: tax distortion is zero when tax rate is zero
- The result due to distortionary intertemporal wedges that grow over time

Model setup

- An economy having a representative consumer who has infinite horizons, that is, choosing C_t and L_t to maximize

$$\sum_{t=0}^{\infty} \beta^t u(C_t, L_t).$$

- The consumer has initial wealth K_0 and earn period-0 labor income of $w_0 L_0$. The consumer saves $w_0 L_0 - C_0$ in the initial period and has a life-time consumer's budget constraint

$$\sum_{t=1}^{\infty} (C_t - w_t L_t) \prod_{s=1}^{t} (1 + r_{s-1})^{-1} \leq K_0 + w_0 L_0 - C_0, \tag{5}$$

where r_t is the after-tax return earned by capital at period t.

- The economy including the government also has an economy-wide budget constraint

$$C_t + G_t + K_{t+1} \leq F_t(K_t, L_t) + K_t,$$

where G_t is exogenous government expenditure.

Solving the model

- Clearly the consumer's budget constraint is binding. FOCs with respect to C_t and L_t results in (denote the Lagrangian multiplier of the consumer's budget constraint as λ)

$$\beta^t \frac{\partial u}{\partial C_t} = \lambda \prod_{s=1}^{t} (1 + r_{s-1})^{-1},$$

and

$$\beta^t \frac{\partial u}{\partial L_t} = -\lambda w_t \prod_{s=1}^{t} (1 + r_{s-1})^{-1}.$$

- We therefore have

$$w_t \frac{\partial u}{\partial C_t} = -\frac{\partial u}{\partial L_t}, \tag{6}$$

and

$$\frac{\partial u}{\partial C_t} = \frac{\partial u}{\partial C_{t+1}} (1 + r_t) \beta \tag{7}$$

Solving the model, continued

- Equation (7) is a recursive expression, going all the way back to date 0 results in

$$\frac{\partial u}{\partial C_0} = \frac{\partial u}{\partial C_n}\beta^n \prod_{i=0}^{n-1}(1+r_i). \tag{8}$$

- From equation (6), $w_t = -\frac{\frac{\partial u}{\partial L_t}}{\frac{\partial u}{\partial C_t}}$. Insert equations (8) and (6) into the consumer's budget constraint, which can be rewritten as

$$C_0 - w_0 L_0 + \sum_{t=1}^{\infty}(C_t - w_t L_t)\prod_{s=0}^{t-1}(1+r_s)^{-1} \leq K_0,$$

results in

$$\sum_{t=0}^{\infty}\beta^t[\frac{\partial u}{\partial C_t}C_t + \frac{\partial u}{\partial L_t}L_t] \leq K_0\frac{\partial u}{\partial C_0}.$$

Solving the model, continued

- Now can rewrite the optimization problem as

$$\max_{\{C_t, L_t, K_t\}} \sum_{t=0}^{\infty} \beta^t u(C_t, L_t) \text{ s.t.}$$

$$\sum_{t=0}^{\infty} \beta^t [\frac{\partial u}{\partial C_t} C_t + \frac{\partial u}{\partial L_t} L_t] \leq K_0 \frac{\partial u}{\partial C_0} (\lambda),$$

  and

$$C_t + G_t + K_{t+1} \leq F_t(K_t, L_t) + K_t \ (\mu_t).$$

- FOC with respect to C_t results in

$$\beta^t \{\frac{\partial u}{\partial C_t} - \lambda [\frac{\partial u}{\partial C_t} + C_t \frac{\partial^2 u}{\partial C_t^2} + \frac{\partial^2 u}{\partial L_t \partial C_t} L_t]\} = \mu_t. \tag{9}$$

- FOC with respect to K_t results in

$$\mu_t (1 + \frac{\partial F}{\partial K_t}) = \mu_{t-1}. \tag{10}$$

Steady state solution

- Steady state implies that $C_t = C$, $L_t = L$ and $K_t = K$ $\forall t$.
- Equation (9) then implies that $\mu_t = \beta\mu_{t-1}$. Insert into equation (10) results in

$$\beta(1 + \frac{\partial F}{\partial K_t}) = 1.$$

- From equation (7), we have

$$\beta(1 + r_t) = 1,$$

  implying that

$$r_t = \frac{\partial F}{\partial K_t},$$

  with the right hand side being the pre-tax return to investors in a competitive market (think of maximization problem of $F(K, L) - rK - wL$). Therefore the optimal tax rate on capital should be zero in the steady state.

More on the zero optimal capital income tax result

- If the government needs to raise revenue, it may need to impose a capital income tax at date 0, followed by zero capital income tax afterwards. However, this may generate "time inconsistency" problems as government may not be able to commit not to impose such capital income tax in subsequent periods.
- Crucial assumption is that consumers live forever; finite lifetimes, however, would still require optimal capital income tax not being too large.
- The result is established for the steady state. Optimal capital income tax during transition periods is still open question (and answers have been provided based on simulations).
- The result is based on a representative consumer. Optimal capital income tax in an economy with heterogenous consumers and different age cohorts may be quite different.