Advanced Public Economics
Professor Panu Poutvaara

8. OPTIMAL WAGE TAXATION

8.1. Introduction

- Wage taxation is used to finance both public consumption and income transfers. It is the most important way for the government to affect income distribution

- In this chapter, the focus is on income redistribution. It is typically assumed that the government transfers the tax revenue it collects from high-income earners to low-income earners. This is without loss of generality. We could interpret that the government collects uniform lump-sum taxes to finance its revenue requirement for other public expenditures.

- Most important contributions to the theory by Sir James Mirrlees (1936-) and Peter Diamond (1940-)

- Mirrlees shared the Nobel Prize with William Vickrey in 1996 “for their fundamental contributions to the economic theory of incentives under asymmetric information.” In the Mirrlees approach, income tax system is viewed as an incentive system that the government poses to its citizens, subject to asymmetric information as the government cannot observe individual abilities. The level of income transfers from the rich to the poor is restricted, even if the government were Rawlsian, by the need to provide the rich incentives not to mimick the poor.

- A classic article: Mirrlees, James A (1971): An Exploration in the Theory of Optimum Income Taxation. Review of Economic Studies 36, 175-208.

8.2. Basic assumptions of the Mirrlees model

- Citizens differ in their productivities

- Earned income depends on hours worked. Working causes a utility cost. Citizens choose how many hours they work.

- Labor market is competitive and the hourly wage rate corresponds to productivity.

- The government cannot observe productivity or hours worked, but only earned total income.

- The government is assumed to be utilitarian and return all tax revenue to citizens

- Tax system can be assumed to be linear or nonlinear

- The government chooses the tax system subject to its own budget constraint, and the incentive compatibility constraints that citizens choose their hours worked optimally, subject to the tax system

- If the tax system is nonlinear, also the household budget constraint is nonlinear. As a consequence, the optimal number of hours worked need not be unique. As a tiebreaking rule, it is assumed that if citizens are indifferent between two different numbers of hours worked, they choose the higher.

8.3. Optimal linear income tax

- n consumers

- Consumer *i* has wage rate w_i. Unobservable to the government

- All consumers have similar utility function $U = U(C_i, L_i)$, where C_i is private consumption and L_i is labor supply. These satisfy:

$U_C > 0$, $U_L < 0$, $U_{CC} < 0$, $U_{LL} \leq 0$, $\lim_{L\to 1} U_L = -\infty$, and $\lim_{L\to 0} U_L = 0$

- Taking into account the income transfer α, consumer *i* faces tax burden $T_i = -\alpha + tw_iL_i$. This can be positive or negative

- Here, *t* is the marginal tax rate. To simplify notation, define $t \equiv 1-\beta$. The consumer's tax burden can now be written as $T_i = -\alpha + (1-\beta)w_iL_i$

- Citizen's budget constraint $C_i = \alpha + \beta w_i L_i$

- Even is the marginal tax rate is constant, the tax system is progressive in that the average tax burden is increasing in earned income. The model corresponds to having a basic income / negative income tax system

- Consumer *i*'s Lagrangean:

$$L = U(C_i, -L_i) + \lambda_i(\alpha + \beta w_i L_i - C_i)$$

- This allows to solve an indirect utility function $V^i = V(\alpha, \beta w_i)$

- By the Envelope Theorem,

$$\frac{\partial V^i}{\partial \alpha} = \lambda_i \quad \text{and} \quad \frac{\partial V^i}{\partial \beta} = \lambda_i w L_i$$

- The government's problem: Max $\sum_i V^i$ with respect to α and β so that the net revenue that the government collects R is zero: $R = \sum_i \left[(1-\beta) w_i L_i - \alpha\right] = 0$

- The government's Lagrangean: $L^G = \sum_i V^i + \mu \sum_i \left[(1-\beta) w_i L_i - \alpha\right]$, where μ is the government's budget constraint's Lagrange multiplier

- The first-order conditions:

α: $\sum_i \lambda_i + \mu(1-\beta)\sum_i w_i \frac{\partial L_i}{\partial \alpha} - \mu n = 0$

β: $\sum_i \lambda_i w_i L_i - \mu \sum_i w_i L_i + \mu(1-\beta)\sum_i w_i \frac{\partial L_i}{\partial \beta} = 0$

μ: $\sum_i \left[(1-\beta) w_i L_i - \alpha\right] = 0$

- Define next gross labor income $z_i \equiv w_i L_i$. Notice that z_i is a function of α, β, and w_i.

- The first two first-order conditions can be written as:

(1) $$\sum_i \lambda_i = \mu n - \mu(1-\beta)\sum_i \frac{\partial z_i}{\partial \alpha}$$

(2) $$\sum_i \lambda_i z_i = \mu \sum_i z_i - \mu(1-\beta)\sum_i \frac{\partial z_i}{\partial \beta}$$

And the government budget constraint can be written as

(3) $$(1-\beta)\sum_i z_i - n\alpha = 0$$

- Divide next equation (2) by (1) on both sides:

(4) $$\frac{\sum_i \lambda_i z_i}{\sum_i \lambda_i} = \frac{\sum_i z_i - (1-\beta)\sum_i \frac{\partial z_i}{\partial \beta}}{n - (1-\beta)\sum_i \frac{\partial z_i}{\partial \alpha}}$$

- Next differentiate equation (3) so that the budget constraint is maintained:

(5) $$\frac{d\alpha}{d\beta}\Big|_{dR=0} = -\frac{\sum_i z_i - (1-\beta)\sum_i \frac{\partial z_i}{\partial \beta}}{n - (1-\beta)\sum_i \frac{\partial z_i}{\partial \alpha}}$$

- The LHS (left-hand side) of equation (4) can be interpreted as welfare weighted average effective labor supply. Welfare weights are the marginal utilities that consumers derive from income. Denote the LHS of (4) by $z(\lambda)$. By equation (5),

$$z(\lambda) = -\frac{d\alpha}{d\beta}\Big|_{dR=0}$$

- Denote the average gross wage income by $\bar{z} = \bar{z}(\alpha, \beta)$. From now on, upper bar is used to denote averages. Equation (4) gives

(6) $$z(\lambda) = \frac{\bar{z} - (1-\beta)\overline{\frac{\partial z}{\partial \beta}}}{1 - (1-\beta)\overline{\frac{\partial z}{\partial \alpha}}}$$

- Totally differentiating $\bar{z} = \bar{z}(\alpha, \beta)$ gives

(7) $$\left.\frac{d\bar{z}}{d\beta}\right|_{dR=0} = \frac{d\bar{z}}{d\beta} + \frac{d\bar{z}}{d\alpha}\left.\frac{d\alpha}{d\beta}\right|_{dR=0} = \frac{d\bar{z}}{d\beta} - \frac{d\bar{z}}{d\alpha} z(\lambda)$$

- Rearrange (6):

$$z(\lambda) - z(\lambda)(1-\beta)\overline{\frac{\partial z}{\partial \alpha}} = \bar{z} - (1-\beta)\overline{\frac{\partial z}{\partial \beta}}$$

This can be rearranged as

(8) $$z(\lambda) - \bar{z} = (1-\beta)\left[z(\lambda)\overline{\frac{\partial z}{\partial \alpha}} - \overline{\frac{\partial z}{\partial \beta}}\right]$$

- Insert (7) into (8):

(9) $$z(\lambda) - \bar{z} = -(1-\beta)\left.\frac{d\bar{z}}{d\beta}\right|_{dR=0}$$

- Equation (9) yields

$$1-\beta=\frac{\bar{z}-z(\lambda)}{\frac{d\bar{z}}{d\beta}|_{dR=0}}$$

- We next return to notaation *t*. As *t=1-β*, $\frac{d\bar{z}}{d\beta}|_{dR=0}=-\frac{d\bar{z}}{dt}|_{dR=0}$. Thus,

(10) $$t=\frac{\bar{z}-z(\lambda)}{-\frac{d\bar{z}}{dt}|_{dR=0}}$$

- Equation (10) defines an optimal tax system implicitly. Explicit solution would require parameterization of the utility function and of the ability distribution.

- Inserting *z(λ)* and $\bar{\lambda}=\frac{\sum_i \lambda_i}{n}$ into the nominator of (10) gives

$$\bar{z}-z(\lambda)=\bar{z}-\frac{\sum_i \lambda_i z_i}{\sum_i \lambda_i}=\bar{z}-\sum_i \frac{\lambda_i}{n\bar{\lambda}}z_i=-Cov\left(\frac{\lambda_i}{\bar{\lambda}},z_i\right).$$

- When the correlation between individual's income and his or her income's social marginal utility is negative, the covariance is negative. So, the nominator of (10) is positive.

- If a tax increase that maintains the government budget constraint reduces gross average income, which is a reasonable assumption as otherwise the government could increase its tax revenue by cutting wage taxes, the denominator of (10) is positive. Thus, t is positive. By the government budget constraint, also α is positive.

- If the downward distortions that wage taxation causes on labor supply increases (meaning that the denominator of (10) increases), the optimal wage tax rate is reduced.

- The analysis can be extended to a general Bergson-Samuelson social welfare function. For a comprehensive analysis, see Gareth Myles: Public Economics, pages 141-145. This is not required reading. It has been shown that the socially optimal marginal tax rate with linear wage taxation is at most the tax rate that would be optimal if the social welfare function is Rawlsian. This, in turn, is that most the tax rate that would maximize the government's tax revenue.

8.4. Optimal nonlinear wage taxation with two types

- Taxation depends on gross income. Same simplifying assumptions as in 8.3 hold, but tax system may now be non-linear

- The utility function can be written as a function of observable arguments c and y: $V(c,y)$

- Type 2 more productive. Thus, the indifference curves of type 2 have a smaller slope in Y,C space that those of type 1. In other words: a smaller increase in consumption suffices to compensate for the disutility of labor supply needed to generate a given income increase for the more productive type.

- Single crossing property is satisfied if at any income and consumption pair, the indifference curve of the high-wage consumer that passes through that point is flatter than the indifference curve of the low-wage consumer passing through the same point:
$$-\frac{V_y^1}{V_c^1} > -\frac{V_y^2}{V_c^2}$$

- Single-crossing property is sufficient to imply that high-skill consumers earn never less than low-skill consumers

- General nonlinear income tax: $T(y)$
- Consumer's budget constraint $c = y - T(y)$
- Consumer maximizes $V[y - T(y), y]$
- Consumer's first-order condition $V_c(1 - T') + V_y = 0$
- The marginal tax rate can be written using the marginal rate of substitution: $MTR(y^i) = T'(y^i) = \frac{V_y^i}{V_c^i} + 1$.
- When the government solves the pairs of consumption and income (c^1, y^1) and (c^2, y^2), it also defines marginal tax rates.
- The government maximizes: Max $W = V^1(c, y) + V^2(c, y)$ subject to the budget constraint $y^1 + y^2 = c^1 + c^2$ taking into account the incentive compatibility constraint that each productivity types must be willing to choose the intended pair of consumption and gross income. This is also called self selection constraint.
  SS(1,2) $V^1(c^1, y^1) \geq V^1(c^2, y^2)$
  SS(2,1) $V^2(c^2, y^2) \geq V^2(c^1, y^1)$

- Constraint *SS(i,j)* tells that type *i* does not want to mimic type *j*

- Focus on income redistribution from the high-productivity type to the low-productivity type. As a result, only SS(2,1) in binding. Intuition in the class.

$$L = V^1(c^1, y^1) + V^2(c^2, y^2)$$
$$+ \lambda\left[V^2(c^2, y^2) - V^2(c^1, y^1)\right]$$
$$+ \gamma(y^1 + y^2 - c^1 - c^2)$$

- The government's first-order conditions:

(1) $V_c^1 - \lambda \hat{V}_c^2 - \gamma = 0$

(2) $V_y^1 - \lambda \hat{V}_y^2 + \gamma = 0$

(3) $(1+\lambda)V_c^2 - \gamma = 0$

(4) $(1+\lambda)V_y^2 + \gamma = 0$

^(hat) refers to the mimicker = type 2 mimicking 1.

- Solve next the marginal tax rates using $MTR(y^i) = T'(y^i) = \dfrac{V_y^i}{V_c^i} + 1$.

- By (4) and (3):

$\dfrac{V_y^2}{V_c^2} = -1$ or $MTR(y^2) = 0$.

- Using (2) and (1) and rearranging yields $\frac{V_y^1}{V_c^1} = \frac{\lambda \hat{V}_c^2}{\gamma}\left[\frac{\hat{V}_y^2}{\hat{V}_c^2} - \frac{V_y^1}{V_c^1}\right] - 1$. Thus, $MTR(y^1) = \frac{\lambda \hat{V}_c^2}{\gamma}\left[\frac{\hat{V}_y^2}{\hat{V}_c^2} - \frac{V_y^1}{V_c^1}\right] > 0$, as the multiplier is positive and so is the term in brackets by the agent monotonicity condition.

- Even if the marginal tax rate of type 2 is zero, his or her average tax rate is positive. Type 1, in contrast, has negative total tax burden.

- Self selection constraint is binding.

- The levels of consumption can be solved using the tax functions.

- Why is the marginal tax rate zero at the top?

- To see this, make a counterassumption that we are in a point in which the marginal tax rate of type 2 is positive. That is, the slope of the budget constraint in the Y,C space is less than one. As a result, wage taxation distorts the labor supply decision of type 2 at the margin. By cutting the marginal tax rate of type 2 to zero with al higher incomes, type 2 could increase his or her utility by working a bit more. This would be a Pareto improvement, proving that the counter-assumption cannot hold in the equilibrium.

- Type 1 faces a positive marginal tax rate as this is needed to make it less attractove for type 2 to mimic types 1.

8.5. Optimal non-linear wage taxation with continuous productivity distribution

- Results reported based on Mirrlees (1971)
- Continuum of tax payers, wage distribution $(\underline{n}, \overline{n})$ with density function $f(n)$
- Consumers maximize $\max_{x,y} u(x,y)$ subject to $x = ny - T(ny)$
- Self-selection constraint for all n and n':
  $u\{ny(n) - T[ny(n)], y(n)\} \geq u\{n' y(n') - T[n' y(n')], y(n')\}$
- The government maximizes $W = \int_{\underline{n}}^{\overline{n}} u(x,y,n) f(n) dn$ subject to the self-selection constraints and the government budger constraint $\int T[z(n)] f(n) dn = R$
- It can be shown that an optimal non-linear wage tax fulfills the following:
  1. $MTR(\overline{n}) = 0$
  2. $MTR(\underline{n}) = 0$, if everyone works in the optimum
  3. $MTR \geq 0$

- As analytical approach does not deliver more results, need for simulation

- In simulations, parameters are given reasonable values and then it is studied how changes in parameters affect income distribution and tax schedule.

- Important aspects to specify:

1. tax revenue relative to total income
2. income distribution in the absence of taxes
3. labor supply behavior
4. social objective function

- Of these, the first three are positive questions. The fourth one is a normative question

- Results:

- A higher social weight on equal income redistribution implies a higher tax rate
- Higher labor supply elasticity implies a lower tax rate
- More income inequality in the absence of taxes increases the socially optimal level of re-distribution
- Marginal tax rates are increasing in the revenue requirement
- The result of a zero marginal tax rate at the top is very local. Marginal tax rates can be quite high just before the top

- No clear results on the shape of teh tax function. Can be highly non-linear
- Traditionally, there has been a consensus in favor of high marginal tax rates at the bottom of the productivity distribution (however, total tax burden there negative, due to transfers)
- No clear results on whether the marginal tax rate should increase in income, as is common
- Recent result has questioned the optimality of high marginal tax rates at the bottom of income distribution
- Intensive and extensive margin
- The United States: *Earned Income Tax Credit*. The United Kingdom: *Working Families Tax Credit*. In both, tax rates are negative at low incomes, in order to encourage labor force participation. After certain level, income subsidies change to taxes
- Saez, Emmanuel (2002): Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses. The Quarterly journal of Economics 117, 1039-1073.
- Saez presents simulations about an optimal system for the United States under alternative assumptions about labor supply elasticity and welfare weights for various groups

8.6. Conclusion

- With optimal linear income tax, the ideological division line is clear: How high the marginal tax rate should be? The level of transfer then determined by the government budget constraint.

- With a non-linear income tax, the decision is more complicated. Arguments in favor of high, low or even negative marginal tax rates for low-income earners.

- Zero marginal tax rate at the top result is more of theoretical interest than of practical importance. Implementing it optimally would require the government to know the ability distribution.

- One argument in favor of a linear tax schedule is uncertainty related to how the economy works.

- In case the income level of the poor would be maintained, a switch to a linear tax schedule would likely increase the tax burden of those of intermediate incomes, and reduce the burden for those with high incomes. The level of redistribution a political decision also in a linear system

- Debate on the optimal tax system continues, both scientifically and in politics.