Public Economics: Tax & Transfer Policies

Final Exam, January 10, 2016 - 2 hours

The exam is 2 hours long and can be done either in French or English. ***No*** *document whatsoever is allowed.*

1 Income taxation (7 points)

We consider an economy made up of individuals who have identical preferences. An individual earns z_i and consumes $c_i = z_i - T(z_i)$ where $T(.)$ is the (possibly nonlinear) income tax.
Suppose that individual i has a utility function of the form:

$$u_i(c, z) = c - \frac{z_i^0}{1 + \frac{1}{e}} \cdot \left(\frac{z}{z_i^0}\right)^{1+\frac{1}{e}}$$

where $e > 0$ is a parameter (the same for all individuals) and z_i^0 is a parameter specific to individual i. Suppose there is a distribution of z with density $f(z) > 0$ over $[0, \infty)$. The total population is normalized to one so that $\int_0^\infty f(z)dz = 1$.

1) What is the economic meaning of e and z_i^0? (0.5 point)

Answer : e is the elasticity of income with respect to the net-of-tax rate. z_i^0 is the potential income, i.e income reported when the marginal tax rate is zero.

2) Consider a linear income tax system $T(z) = -R + \tau \cdot z$ where $R > 0$ is the demogrant and τ is a flat tax rate. Solve the individual maximisation problem and show that individual i earns $z_i = z_i^0 \cdot (1 - \tau)^e$ when the tax rate is τ. (1 point)

Answer : Program of the individual i :

$$Max_z \; R + (1 - \tau)z_i - \frac{z_i^0}{1 + \frac{1}{e}} \cdot \left(\frac{z_i}{z_i^0}\right)^{1+\frac{1}{e}}$$

$$\text{FOC} : (1-\tau) = \left(\frac{z_i}{z_0}\right)^{1/e}$$

$$\text{FOC} : z_i = z_i^0 \cdot (1-\tau)^e$$

3) Suppose taxes collected are all rebated through the demogrant so that $R = \tau Z$ where Z is average earnings. Solve for the Rawlsian optimal tax rate τ (i.e, the tax rate that maximises the utility of the worst-off individual). Solve for the utilitarian optimal tax rate τ (i.e, the tax rate that maximises the sum of utilities). In both cases, explain the intuition behind the results. (1.5 points)

Answer : The social welfare function is given by:

$$SWF = \int_i f(z_i) u_i((1-\tau)z_i + R, z_i) dz_i$$

$$\text{s.t } R = \tau Z$$

where Z is the average earnings depending on the net of tax rate $1-\tau$.

a) *Rawlsian social welfare function:*
Maximise utility of individuals with $z_i = 0 \Rightarrow u = R = \tau Z$ so Rawlsian optimal rate maximizes tax revenue (to maximize R), i.e τZ, and the optimal tax rate is set to $\tau^\star = 1/(1+e)$. (FOC gives $Z - \tau \frac{dZ}{d(1-\tau)} = 0 \Rightarrow (1-\tau) - \tau e = 0$. In absence of behavioral response, the optimal tax rate would be 100%. As 100% taxation would lead everybody to completely stop working, the optimal tax rate is lower depending on the magnitude of the behavioral responses (e).

b) *Utilitarian social welfare function:*

$$SWF = \int_i f(z_i) u_i((1-\tau)z_i + R, z_i) dz_i$$

$$\Rightarrow SWF = \int_i f(z_i) \left[\frac{(1-\tau)^{1+e} z_i^0}{1+e} + \tau Z\right] dz_i$$

$$\frac{dSWF}{d\tau} = 0 \Rightarrow \int_i f(z_i) \left[-z_i + Z - \tau \frac{dZ}{d(1-\tau)}\right] dz_i = 0$$

$$\Rightarrow Z - Z - \tau \frac{dZ}{d(1-\tau)} = 0 \Rightarrow \tau^\star = 0$$

Given that all utilities are linear, there is no concern for redistribution and hence the optimal utilitarian tax rate is zero.

4) The government asks you to estimate e using two cross-section random sample of individual earnings (this is not panel data) for two consecutive years: year 1 and year 2. In year 1, the tax rate is τ_1. In year 2, the tax rate increase to level τ_2. How would you proceed to estimate e from this data? Provide the regression specification that would allow you to estimate e. State clearly what assumptions would be needed to estimate e without bias. (1.5 points)

Answer : OLS regression $log(z_{it}) = \alpha + e \cdot log(1-\tau_t) + \epsilon_{it}$ pooling together the two cross-sections. $\hat{e}$ is unbiased if absent the tax change, the average log-income would have stayed constant, from year 1 to year 2, ie, $E(\log z_i^0)$ is the same in year 1 and year 2.

5) Let us now assume that the government sets a two bracket tax schedule with a zero marginal tax rate for incomes below $z^\star$ and a marginal tax rate for incomes above $z^\star$ in year 1. The government increases the marginal tax rate above $z^\star$ from τ_1 in year 1 to τ_2 in year 2 (the tax rate below $z^\star$ remains at zero). You have access to panel earnings data following the same n individuals in year 1 and year 2. Explain how you could identify e exploiting this reform and doing a difference-in-difference analysis. Provide the regression specification that would allow you to estimate e. What are the key identification assumptions needed? Discuss the potential biases that arise when applying this method to real world data. (1.5 points)

Answer : Define the treatment group T as individual above $z^\star$ in year 1 and the control group C as individuals below $z^\star$ in year 1. Then compare the log earnings increase for both groups from year 1 to year 2 following the change in log of net-of-marginal-tax-rates. This can be obtained by 2SLS regression:

$$\log(\frac{z_{i2}}{z_{i1}}) = \alpha + e \log(\frac{1-\tau_i 2}{1-\tau_i 1}) + \epsilon_i$$

using the dummy for being in the treatment group as the instrument. $\hat{e}$ is unbiased if absent the tax change, the average log-income change would have stayed the same from year 1 to year 2. In practice, mean reversion, secular growth and non-tax related changes in inequality (differential trends between earnings of control and treated groups) would create a bias.

6) In a very influential paper, Feldstein (1999) stated that the elasticity of taxable income was a sufficient statistic to estimate the efficiency costs of taxation. Is it true? What are the strong assumptions required to apply it? (1 point)

Answer : The assumption is that a reduction in reported incomes due to a tax rate increase has no other effect on tax revenue. Problematic if the reduction in reported incomes is due in part to a shift away from taxable individual income toward other forms of taxable income such as corporate income, or deferred compensation that will be taxable to the individual at a later date (Tax base shifting or timing responses). Problematic if short-term tax responses are larger than longer-term responses. Problematic in presence of externalities (charitable giving or deductions from taxable income).

2 Public goods (5 points)

Consider N identical consumers indexed by $i = 1, ..., N$ with the same utility function:

$$U_i = \log(x_i) + \log(G)$$

where x_i is the consumption of a private good by individual i, and G is a pure public good. For simplicity, we assume that individual earnings and the price of the private good are equal to 1 such that each consumer's budget constraint can be written as:

$$x_i + g_i \leq 1$$

where g_i is the contribution to the public good of individuals i. Total available quantity of the public good is the sum of individual contributions, i.e. $G = \sum_{i=1}^{N} g_i$.

1) What are the two properties of a public good and explain why it can cause a market failure. (1 point)

Answer :

- Non rival in consumption: an individual's consumption of a good does not affect another's opportunity to consume the good
- Non excludable: There is no way to deny someone the opportunity to consume the good

Problem of free-rider: Inefficient private provision

2) Calculate G^d, the equilibrium public good provision when individuals take decentralized decisions. (1.5 points)

Answer: Program of the individual i :

$$Max\ U_i = \log(1 - g_i) + \log(g_i + \sum_{j=1, j\neq i}^{N} g_j)$$

FOC for g_i gives:

$$\frac{\partial U_i}{\partial g_i} = -\frac{1}{1 - g_i} + \frac{1}{g_i + \sum_{j=1, j\neq i}^{N} g_j} = 0$$

All individuals are identicial $\Rightarrow g_i = g and g = \frac{1}{N+1} \Rightarrow G^d = \frac{N}{N+1}$

3) Calculate G^o, the optimum public good provision when a social planner chooses the level of public good such has each individual contributes equally and the following social welfare function is maximized (1.5 points) :

$$SWF = \sum_{i=1}^{N} U_i$$

Answer: Program of the social planner :

$$\text{Max} \sum_{i=1}^{N} \log(1 - \frac{G}{N}) + \log(G) = \text{Max}\ N\left(\log(1 - \frac{G}{N}) + \log(G)\right)$$

FOC for G gives:

$$\frac{\partial SWF}{\partial G} = N\left(-\frac{1}{1-\frac{G}{N}}\frac{1}{N}+\frac{1}{G}\right)=0$$

$$\Rightarrow G^o = \frac{N}{2}$$

4) Without any computations, explain how could the optimum public good provision be obtained through a decentralized mechanism? (1 point)

Answer: Lindhal pricing: individuals should pay different prices for the public good, equal to their own marginal benefit. The idea is to ask each individual i to pay a share τ_i of the public good and too pick a level of G. But incentive to under-report ...

3 Questions (8 points)

1) Atkison and Stiglitz (1980) argues that the person who effectively pays a tax is not necessarily the person upon whom the tax is levied. Without any computations, show graphically the impact of the introduction of a Value Added tax rate τ on the consumer and producer surplus. (1 points)

Answer: See graph on the incidence of the VAT in Lecture 4.

2) Apart from econ grad students, individuals tend to free-ride much less than the theory predicts in public good games. How can you explain that? (1 points)

Answer: Possible explanations: non-monetary pleasure from cooperative outcomes and social norms. Miguel and Gugerty (2005) highlight the role of social sanctions in explaining why ethnic diversity is associated with lower levels of public good provision in Western rural Kenya.

3) If society cares about redistribution, it is desirable to have lower VAT rates on goods that are disproportionately consumed by low income families? (1 point)

Answer: No, because rich people do spend on these goods. Better job to redistribute through income tax + administration issues + lobbying on which goods should have reduced rates.

4) Evaluate the following claims by determining whether each claim is true or false and present a concise explanation for your answer:

i) The empirical observation that those receiving unemployment insurance (UI) benefits remain unemployed longer than those not receiving UI benefits, conditional on unemployment, indicates that UI causes longer unemployment spells. (0.75 point)

ii) Assuming that UI causes longer unemployment spells, this clearly indicates that generosity of the program should be reduced. (0.75 point)

4.i) Answer: False. We need to know why some individuals receive UI and others do not. If the provision of UI benefits is random, then the claim would be true. The problem is that replacement rates are typically a function of past earnings, which are themselves correlated with re-employment prospects.

4.ii) Answer: Answer: False. The downside of the moral hazard created by UI must be weighed against the consumption smoothing benefits of the program.

5) The traditional approach to measuring the impact of the generosity of disability insurance on labor force participation has been to estimate the following linear probability model through OLS, using cross-sectional data on men aged 45-64:

$$LFP_i = \alpha + \beta \cdot RR_i + \gamma \cdot X_i + \varepsilon_i$$

Where LFP_i is a dummy variable equal to 1 if individual i is in the labor force and zero otherwise, RR_i is the potential disability insurance replacement rate that the individual faced while working and X_i is a vector of covariates such as age, region of residence, etc.

i) How do you interpret β? What is its expected sign? (0.5 point)

ii) In the context of disability insurance, why are replacement rates usually larger for low-wage workers than for high-wage workers? How might this bias the estimation of β in the above specification? (0.5 point)

iii) What estimation strategy would you advocate to recover a causal estimate of the generosity of disability insurance on labor force participation? (0.5 point)

5.i) Answer: a 1 percentage point increase in the replacement rate reduces the labor force participation of men aged 55-65 by $-\hat{\beta}$ percentage points. Note that in this specification, neither the dependent variable nor the replacement rates are taken in logs, so the coefficient cannot be interpreted as an elasticity of labor force participation with respect to the generosity of disability insurance. Since more generous disability insurance should increase the incentives to drop out of the labor force, we expect $\hat{\beta}$ to have a negative sign.

5.ii) Answer: disability insurance benefits are usually capped at a certain fixed amount, which implies that replacement rates are a decreasing function of previous wages. Due to omitted variable bias, this phenomenon could lead to wrong inferences as to the impact of disability insurance generosity on labor force participation. The finding that workers with higher potential disability insurance replacement rates are more likely to leave their jobs may indeed simply reflect the fact that low-wage workers have worse employment prospects than higher wage workers. In this case, the coefficient $\hat{\beta}$ will then tend to overestimate the negative impact of disability insurance generosity on labor force participation.

5.iii) Answer: what is needed to identify the behavioral impact of disability insurance is variation in program generosity, which is independent of underlying tastes for work. This variation can be provided, for instance, by the increase in benefits in some states or regions relatively to other states or regions. This type of variation is used by Gruber (2000) who exploits the fact that the level of disability benefits remained roughly constant in Canada during the 1980s whereas the rest of Canada raised its benefits levels by 36% in 1987. In this empirical setting, the causal impact of the generosity of disability insurance on labor force participation can be estimated using a difference-in-differences approach.

6) In the presence of uncertainty about the marginal cost of pollution reduction, which policy instrument should the government favor to control pollution? (1 point)

Answer: Command-and-control instruments: set standards on the level of pollution, and controls to monitor and enforce the standard (Restriction on emissions, Mandates for the adoption of specific technologies). Market-based instruments: rely on market or price mechanisms to give incentives to polluters to reduce emissions (Pigouvian taxes or subsidies, Tradeable permits).

7) Some economists have discussed the possibility of taxation based on height. Explain the rationale for such a tax and the issues related to its implementation. (1 point)

Taxation based on individual characteristics that can be observed by the government, are immuable and are correlated with endowments and ability can be optimal for redistribution as it does generetae any behavioral responses. (Tagging, Akerlof 1978). Following that, Mankiw and Weinzierl (2010) proposes to tax individuals according to their height. Problem: Perverse incentives by stigmatisation of tagged individuals, Horizontal equity issue.