Answers

Final exam in Public Finance - Spring 2018

3-hour closed book exam

By
Jakob Egholt Søgaard

Part 1: Social welfare and redistribution

Consider an economy consisting of two individuals (low skilled (L) and high skilled (H)). Both individuals maximize the following utility function

$$U_i(c_i, h_i) = u(c_i) - v(h_i), \tag{1}$$

subject to the budget constraint

$$c_i = w_i h_i - T_i, \tag{2}$$

where c_i is consumption, h_i is labor supply, T_i is an individual lump sum tax and w_i is the individual wage rate where $w_H > w_L$.

(1A) *Show that the individuals' utility maximization implies the following first order condition*

$$w_i u'(c_i) = v'(h_i) \tag{3}$$

#

The individual utility maximization problem can be solved by substituting equation (2) into equation (1) and differentiating wrt. h_i:

$$U_i(c_i, h_i) = u(c_i) - v(h_i) = u(w_i h_i - T_i) - v(h_i)$$

$$\Rightarrow \frac{\partial U_i(c_i, h_i)}{\partial h_i} = 0 \Leftrightarrow w_i u'(c_i) = v'(h_i),$$

which was what we were asked to show.

The government sets lump sum taxes in order to maximize social welfare given by

$$W = \sum_i U_i(c_i, h_i) \tag{4}$$

subject to the budget constraint

$$T_L + (1-q)T_H = 0, \tag{5}$$

where $0 \le q \le 1$ is a revenue loss when collecting taxes from the high skilled.

(1B) *Show that at the social optimum, the government will set lump taxes such that $(1-q)u'(c_L) = u'(c_H)$. Discuss how q affects the amount of redistribution and the implications for the utility levels of the two individuals.*

#

To find the social optimum, we first substitute the equations (1), (2) and (5) into equation (4)

$$\begin{aligned} W &= u(w_H h_H - T_H) - v(h_H) + u(w_L h_L - T_L) - v(h_L) \\ &= u(w_H h_H - T_H) - v(h_H) + u(w_L h_L + (1-q)T_H) - v(h_L), \end{aligned}$$

and differentiate wrt. T_H

$$\begin{aligned} \frac{\partial W}{\partial T_H} &= -u'(w_H h_H - T_H) + \underbrace{w_H u'(w_H h_H - T_H)\frac{\partial h_H}{\partial T_H} - v'(h_H)\frac{\partial h_H}{\partial T_H}}_{(A)} \\ &\quad + (1-q)u'(w_L h_L + (1-q)T_H) + \underbrace{w_L u'(w_L h_L + (1-q)T_H)\frac{\partial h_L}{\partial T_H} - v'(h_L)\frac{\partial h_L}{\partial T_H}}_{(B)} = 0. \end{aligned}$$

Using the individuals' first order condition in equation (3) we see that (A) and (B) in the above equation is 0 (the envelope theorem), and we therefore get the following condition for the social optimum

$$-u'(w_H h_H - T_H) + (1-q)u'(w_L h_L + (1-q)T_H) \Leftrightarrow (1-q)u'(c_L) = u'(c_H),$$

which was what we were asked to show.

From this equation we see that the level of redistribution depends negatively on the cost of redistribution (q). Starting from $q = 0$, we have $u'(c_L) = u'(c_H) \Leftrightarrow c_L = c_H$. That is, when it is costless to redistribution, the government will set taxes such that both individuals have the same consumption level.

From the individuals' first order condition we see that the high skilled individual in this case will choice to work more than the low skilled individual (when $u'(c_L) = u'(c_H)$ then $w_H > w_L \Leftrightarrow v'(h_H) > v'(h_L) \Leftrightarrow h_H > h_L$). With $q = 0$ the high skilled individual will therefore have the same level of consumption as the low skilled, but at the same time work more, which implies that the high skilled individual will end up with lower utility than the low skilled.

With a higher q, the government will set taxes such that $(1-q)u'(c_L) = u'(c_H) \Rightarrow u'(c_L) > u'(c_H) \Leftrightarrow c_L < c_H$, which implies a lower level of redistribution and - as a consequence - a lower level of utility for the low skilled individual and higher utility for the highskilled individual.

(1C) *How does the level of redistribution in (1B) change if the goverment weights the utility of the low skilled individual higher than the utility of the high skilled. That is, if the social welfare function is given by $W = gU_L(c_L, h_L) + U_H(c_H, h_H)$ with $g > 1$?*

#

Adding the social welfare weight g social welfare function and repeating the calculations in the answer to question (1B) gives the following condition for the social optimum

$$g(1-q)u'(c_L) = u'(c_H).$$

From this equation we see that the level of redistribution depends positively on the welfare weight on the low skilled (g). For example, with $q = 0$ and $g > 1$ we have $gu'(c_L) = u'(c_H) \Rightarrow u'(c_L) < u'(c_H) \Leftrightarrow c_L > c_H$. Likewise for any $q > 0$, a social welfare weight $g > 1$ implies more redistribution compare to the situation in (1B).

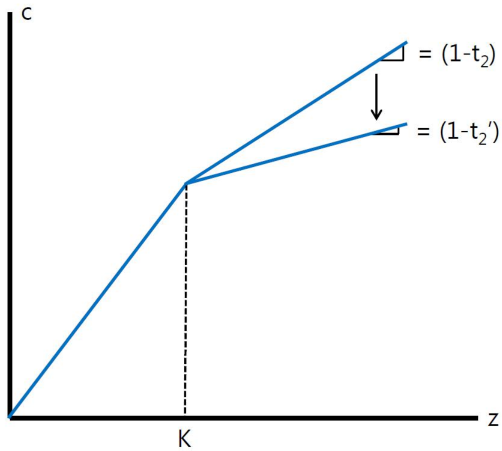
Part 2: Taxation of high income earners

Consider a large number of individuals who face a two-bracket tax system with a marginal tax rate t_1 on pre-tax income (z) below a threshold K and a marginal tax rate $t_2 > t_1$ on the income exceeding K.

(2A) *Illustrate in a diagram with pre-tax income (z) on the x-axis and after-tax income (c) on the y-axis, the budget set create by the tax system, and show how the budget set changes, when the top tax rate (t_2) is increased. Discuss how the labor supply of the individuals are affected by income and substitution effects depending on their initial pre-tax income.*

#

The budget set create by the tax system is illustrated in the figure below.



The slope of the budget set is given by 1 - the marginal tax rate and an increase in the top tax rate t_2 therefore decreases the slope of the budget set above the threshold K as illustrated in the figure above.

As the budget set only changes above K, only individuals with an initial income above K are affected by the tax change.

The labor supply of the individuals above K are affected by both income and substitution effects, but the relative strength of the income and substitution effect depends on the distrance between the individuals' initail income and K. One way to see this is with the following argument:

- The strengh of the substitution effect depends on change in marginal tax rate. This change is the same for all individuals above K, and the strengh of the substitution effect is therefore the same for all individuals above K (for a given substitution elasticity).

- The strengh of the income effect depends on the change in disposable income caused by the tax change (holding fixed pre-tax income). For individuals with an initial pre-tax income just above K, this change is small, and for these individuals the income effect is therefore small. For individiuals with an initial pre-tax income further above K, the change in disposable income is larger and therefore also the income effect.

The above effects may also be compared to a general tax increase (in a proportional tax system where $t_1 = t_2 = t$), where it can be noted that the substitution effect is the same in the two cases (for the affected part of the population), while the income effect is smaller in the case of an increase in top tax. Finally, it can also be noted that the kink point in the budget set at K creates "bunching".

The revenue from the top tax is given by:

$$R = t_2(\bar{z} - K)N, \tag{6}$$

where $\bar{z}$ is the average pre-tax income for the individuals above the threshold K and N is the number of top tax payers. Assume that $\bar{z}$ depend positively on the after-tax rate $(1 - t_2)$ with a constant elasticity ε.

(2B) *Show that the effect of a marginal increase in t_2 on the government's revenue can be written as:*

$$\frac{dR}{dt_2} = \left(\frac{1}{\alpha} - \frac{t_2}{1 - t_2}\varepsilon\right) N\bar{z}, \quad \textit{where } \alpha = \frac{\bar{z}}{\bar{z} - K}. \tag{7}$$

Provide an interpretation for α and comment on how $\frac{dR}{dt_2}$ depend on α.
#

Differentiating equation (6) wrt. t_2 we obtain

$$\frac{dR}{dt_2} = \underbrace{(\bar{z} - K)N}_{\textit{Mechanical effect}} + \underbrace{t_2 \frac{d\bar{z}}{dt_2} N}_{\textit{Behavioral effect}}$$

$$= \left(\frac{\bar{z} - K}{\bar{z}} + \frac{t_2}{1 - t_2} \frac{d\bar{z}}{d(1 - t_2)} \frac{d(1 - t_2)}{dt_2} \frac{1 - t_2}{\bar{z}}\right) N\bar{z},$$

where we used the chain rule to write $\frac{d\bar{z}}{dt_2}$ as $\frac{d\bar{z}}{d(1-t_2)} \frac{d(1-t_2)}{dt_2}$. Using that $\frac{d(1-t_2)}{dt_2} = -1$ along with the definition of the labor supply elasticity $\varepsilon = \frac{d\bar{z}}{d(1-t_2)} \frac{1-t_2}{\bar{z}}$, we obtain

$$\frac{dR}{dt_2} = \left(\frac{\bar{z} - K}{\bar{z}} - \frac{t_2}{1 - t_2}\varepsilon\right) N\bar{z} = \left(\frac{1}{\alpha} - \frac{t_2}{1 - t_2}\varepsilon\right) N\bar{z},$$

which is what we were asked to show.

We see from the equation that the effect of an increase in t_2 on the government revenue depend negatively on α, which is a measure of the income constration of the top tax payers. When the average income of the top tax payers ($\bar{z}$) is close to the top tax threshold (K), α is high and vice versa. As only the income above K is taxed at the top tax rate (t_2), a high α implies that only a small portion of the top tax payers' income is affected by the higher tax rate and the mechanical revenue effect is therefore small. The behavioral effect is - on the other hand - unaffected by the α and a higher α hence reduces the total effect on the government revenue.

(2C) *Show that the revenue maximizing top tax rate ($\hat{t}$) is given by:*

$$\hat{t} = \frac{1}{1+\varepsilon\alpha} \tag{8}$$

Comment on the expression and describe how it differs from the revenue maximizing tax rate in a tax system with a constant marginal tax rate on all income (a proportional tax system). #

We find the revenue maximizing top tax rate by setting (7) equal to zero

$$\frac{dR}{dt_2} = 0 \Leftrightarrow= \left(\frac{1}{\alpha} - \frac{t_2}{1-t_2}\varepsilon\right) N\bar{z} = 0$$

$$\Leftrightarrow 1 = \frac{t_2}{1-t_2}\varepsilon\alpha \Leftrightarrow 1 - t_2 = t_2\varepsilon\alpha \Leftrightarrow 1 = (1+\varepsilon\alpha)\, t_2$$

$$\Leftrightarrow t_2 = \frac{1}{1+\varepsilon\alpha} = \hat{t}.$$

We see from the equation that the revenue maximizing top tax rate is decreasing in both the labor supply elasticity and α. The intuition for the effect of α is the same as in the answer to (2B). That is, a higher α reduces the mechanical effect of the tax increase, because only a small portion of the top tax payers' income exceed the top tax cut-off K. In contrast, the intuition for the negative effect of ε comes from the fact that a higher labor supply elasticity increase the (negative) behavioral effect. Both a higher α and a higher ε will therefore lower the revenue maximizing top tax rate.

A linear/proportional tax system is characterized by having the same (marginal) tax rate on all income. This is equivalent to having $K = 0$ in the model above. In this case we would there have the following revenue maximizing tax rate

$$\hat{t} = \frac{1}{1+\varepsilon}.$$

As $\alpha \geq 1$, we see that revenue maximizing tax rate in a linear/proportional tax system is always larger than the revenue maximizing top tax rate in a progressive tax system.

Part 3: Tax incidence and empirical measurement

Consider a perfectly competitive labor market with many firms who demand labor and many workers who supply labor. The government levies a unit tax (t) on labor, so that the cost per *unit of labor for firms is given by $w_F = w_W + t$, where w_W is the wage rate paid to workers. Given a marginal increase in t, the share of the extra tax burden born by workers (I_W) and firms (I_F), respectively, may approximately be written as*

$$I_W \approx \frac{\varepsilon_F}{\varepsilon_F + \varepsilon_W}, \quad I_F \approx \frac{\varepsilon_W}{\varepsilon_F + \varepsilon_W}, \tag{9}$$

where ε_W is the elasticity of labor supply with respect to wage rate w_W, while ε_F is the (numerical) elasticity of labor demand with respect to the wage cost w_F.

(3A) *Describe how the economic incidence depends on the elasticities and the economic intuition behind these relationships.*

#

The economic incidence is determined by the relative size of the elasticities. The part with the lowest elasticity will bear most of the economic burden. Consider, for example, the case where labor supply is completely inelastic ($\varepsilon_W = 0$). In this case, workers supply a fixed number of work hours, completely independent of the af ter tax wage rate. A higher tax on firms will reduce labor demand but, because the labor supply curve is vertical, this will go directly into a lower wage of workers. In the other extreme case, labor supply is perfectly elastic ($\varepsilon_W \to \infty$) implying that workers are will to work any number of hours as long as the wage rate is above some fixed level. In this case, firms will reduce demand until the point where they are willing to pay workers the original wage plus the tax, and firms will therefore bear the full burden of the tax. These points may be illustrated graphically.

The article "The Incidence of Mandated Maternity Benefitts" in the American Economic Review (1994) by Jonathan Gruber studies the incidence of mandated maternity benefits paid by employers through health insurances for their employees. Below (next page) is a copy of Table 3 from the article.

(3B) *Describe the empirical analysis and explain, using Table 3 in Gruber (1994), how the author arrives at his estimate.*

#

The empirical analysis in Gruber (1994) shows that the extra health insurance costs of the firms due to mandated maternity benefits are shifted to the employees through lower wages. According to the estimate in Table 3, hourly wage rates decrease on average by 5.4 percent of those affected.

The empirical analysis exploits that some states passed laws prohibiting treating pregnancy different from other illness and thereby imposing higher health insurance costs of firms. This creates exogenous variation in labor costs of fertile women compared to other groups in the states passing the laws (within state variation) and also exogenous variation in labor costs of fertile women living in the states passing the laws compared to fertile women living in other st ates (across state variation).

Gruber constructs a DiDiD estimator that exploits the exogenous variation in both dimensions at the same time. He first compute the before-after reform difference in wages of fertile women in the treatment states relative to the control states and he then subtracts the same DiD computed for other individuals not subject to the reform (and arrive thereby at the result of 5.4 percent reported in the bottom of the table). This method controls for state-specific time trends (within-state difference) and group-specific time trends (between-state difference).

(3C) *What is the main identifying assumption needed in (3B) for the estimate to be the causal effect of mandated maternity benefits on the hourly wages of fertile women? Describe how you could validate the main identifying assumption and what kind of data you would need to do so.*

#

The identifying assumption is identical trend-differences between fertile women and others in treatment states and control states without the reform. That is, absent the reform the difference between the change in hourly wages for fertile women and others in the treated states should been the same as in the control states. This is a weaker assumption than the common trend assumption underlying the standard DiD identification strategy and is possible because of the two dimensions of variation in the data.

It is not possible to prove whether or not this assumption is fullfilled, but the assumption can be validated with various placebo tests. For example by repeating the same empirical analysis as in Table 3 over a period with no reforms. In this case we should not expect any trend differences between fertile women and others in treatment states and control states, and a DiDiD estimate of zero would therefore validate the empirical results in Gruber (1994).

Imagine that employers shifted the costs of the mandated maternity benefits to all workers within a given state and not just to fertile women.

(3D) *How would this effect be captured in Gruber's empirical analysis in Table 3? That is, what would you expect the DDD estimate to be in this case?*

#

If employers shift the costs of the mandated maternity benefits to all workers within a given state, then average hourly wages will be lower for both fertile women and others in the treated states. The effect of the mandated maternity benefits would therefore be captured by the state-specific time trends, and we should therefore expect the DDD estimate to be zero.

TABLE 3—DDD ESTIMATES OF THE IMPACT OF STATE MANDATES ON HOURLY WAGES

Location/year	Before law change	After law change	Time difference for location
A. *Treatment Individuals: Married Women, 20–40 Years Old:*			
Experimental states	1.547 (0.012) [1,400]	1.513 (0.012) [1,496]	−0.034 (0.017)
Nonexperimental states	1.369 (0.010) [1,480]	1.397 (0.010) [1,640]	0.028 (0.014)
Location difference at a point in time:	0.178 (0.016)	0.116 (0.015)	
Difference-in-difference:	−0.062 (0.022)		
B. *Control Group: Over 40 and Single Males 20–40:*			
Experimental states	1.759 (0.007) [5,624]	1.748 (0.007) [5,407]	−0.011 (0.010)
Nonexperimental states	1.630 (0.007) [4,959]	1.627 (0.007) [4,928]	−0.003 (0.010)
Location difference at a point in time:	0.129 (0.010)	0.121 (0.010)	
Difference-in-difference:	−0.008: (0.014)		
DDD:	**−0.054** **(0.026)**		

Notes: Cells contain mean log hourly wage for the group identified. Standard errors are given in parentheses; sample sizes are given in square brackets. Years before/after law change, and experimental/nonexperimental states, are defined in the text. Difference-in-difference-in-difference (DDD) is the difference-in-difference from the upper panel minus that in the lower panel.