Economics 2450A: Public Economics
Section 7: Optimal Top Income Taxation

Matteo Paradisi

October 24, 2016

In this Section we study the optimal design of top income taxes.[1] We have already covered optimal top income taxation in a simple Mirrlees framework in Section 2. Today, we will start from the "trickle down" model with endogenous wages introduced by Stiglitz (1982). We then analyze an example of optimal taxation in a general equilibrium model where workers choose between different occupations/sectors (Rothschild Scheuer, 2016). Finally, we present a model where top earners respond to taxes on three margins: labor supply, tax avoidance, and compensation bargaining (Piketty Saez Stantcheva, 2014).

1 Trickle Down: A Model With Endogenous Wages

Stiglitz (1982) studies a model with two unobservable types and endogenous wages. Suppose there are two types of workers: H (high skill) and L (low skill). For simplicity we assume they have equal mass and work l_i hours. The utility for a generic type i is $u(c_i, l_i)$. Work is the only input in the constant return to scale (CRS) production function $F(l_L, l_H)$. With competitive labor markets wages are equal to the marginal product of labor:

$$w_i = \frac{\partial F(l_L, l_H)}{\partial l_i} \tag{1}$$

The standard Mirrlees model implicitly assumes a linear production $F(l_L, l_H) = \theta_L l_L + \theta_H l_H$ where θ_i is the ability of agent i so that wages are $w_i = \theta_i$.

The resource constraint of this economy is:

$$\sum_i c_i \leq F(l_L, l_H)$$

The government assigns linear welfare weights ψ_L and ψ_H to the two types. If $\psi_H < \psi_L$ the government wants to redistribute to low types and we know that in equilibrium the incentive constraint for the high type is binding:

$$u(c_H, l_H) = u\left(c_L, \frac{w_L l_L}{w_H}\right) \tag{2}$$

We solve the problem with the following Lagrangian:

$$\begin{aligned} L \quad = \quad & \psi_L u(c_L, l_L) + \psi_H u(c_H, l_H) + \lambda \left[F(l_L, l_H) - \sum_i c_i\right] \\ & + \mu \left[u(c_H, l_H) - u\left(c_L, \frac{w_L l_L}{w_H}\right)\right] + \sum_i \eta_i \left(w_i - F_i(l_L, l_H)\right) \end{aligned} \tag{3}$$

[1] The first two paragraphs of this section are based on notes by Florian Scheuer.

λ is the marginal value of public funds, μ is the value of relaxing the incentive constraint for type H and ηs are the multipliers on the constraints in (1).

We derive the optimal marginal tax rate for the high type by optimally choosing c_H and l_H. The FOCs are respectively:

$$[\psi_H + \mu]\, u_c\,(c_H, l_H) = \lambda \tag{4}$$

$$[\psi_H + \mu]\, u_l\,(c_H, l_H) = -\lambda F_H\,(l_H, l_L) + \sum_i \eta_i F_{iH}\,(l_L, l_H) \tag{5}$$

The optimal labor supply choice implies the following labor wedge:

$$T'\,(z_H) = 1 + \frac{u_l\,(c_H, l_H)}{u_c\,(c_H, l_H)\, w_H}$$

Using (4) and (5) we can rewrite the labor wedge as follows:

$$T'\,(z_H) = 1 + \frac{-\lambda F_H\,(l_H, l_L) + \sum_i \eta_i F_{iH}\,(l_L, l_H)}{\lambda F_H\,(l_L, l_H)} = \frac{\sum_i \eta_i F_{iH}\,(l_L, l_H)}{\lambda F_H\,(l_L, l_H)} \tag{6}$$

The sign of (6) depends on η_L and η_H. In order to sign them, we exploit the government optimal choice of w_i characterized by the following FOCs:

$$-\mu u_l \left(c_L, \frac{w_L l_L}{w_H}\right) \frac{l_L}{w_H} + \eta_L = 0$$

$$\mu u_l \left(c_L, \frac{w_L l_L}{w_H}\right) \frac{w_L l_L}{w_H^2} + \eta_H = 0$$

Since $u_l < 0$, they imply $\eta_L < 0$ and $\eta_H > 0$. The CRS technology and concavity imply $F_{HL} > 0$ and $F_{HH} < 0$, which means complementarity between the two factors of production. Therefore, $\sum_i \eta_i F_{iH}\,(l_L, l_H) < 0$ and we conclude that $T'\,(z_H) < 0$. Top earners are subsidized at the margin because their labor raises the wages of lower earners. By closing the gap between the two wages the government can relax the incentive constraint for the high type and allow for additional redistribution. The result is entirely driven by the complementarity of the two factors in the production function, which generates a "positive externality" of the high type on the low type. In the classical Mirrlees model with linear technology there is no complementarity and top incomes are not subsidized.

2 Taxation in the Roy Model and Rent-Seeking

In this paragraph we study a more general model introduced by Rothschild and Scheuer (2013, 2016) where individuals can choose the sector where they work, how much they work and have a multidimensional vector of skills (one for each sector). We present a simple example with two activities and a two dimensional skill vector. Workers can choose between two activities: a traditional productive activity where the wage reflects the social marginal product of labor and rent-seeking where the marginal product of labor is zero and workers compete for a fixed rent μsuch that wages are proportional to μ/E, with E being the total effort in the rent-seeking sector. Every individual has a skill vector (θ, φ) such that θ is the ability in the productive sector and φ is ability in the rent-seeking sector. Suppose there are only two types of workers in the economy: productive workers with $\theta = \varphi = 1$ and rent-seekers with $\theta = 0$ and $\varphi = \varphi_R$.

The total rent-seeking effort is:

$$E = \varphi_R e_R + \lambda_P e_P$$

where λ_P is the fraction of productive workers working in the rent-seeking sector. Productive workers are indifferent between the two sectors when the wage in the rent-seeking sector is equal to 1 (the marginal product of labor in productive sector) and we have $\mu/E = 1$ implying $\mu = E$. If instead $E > \mu$ they would all work in the traditional sector; while when $E < \mu$ they would all choose the rent-seeking activity.

Suppose that preferences are quasi-linear $u(c, e) = c - h(e)$. It can be shown that the optimal allocation involves an interior equilibrium where productive workers are indifferent between the two occupations. If λ_P is the share of productive workers working in the rent-seeking sector, we have:

$$E = \varphi_R e_R + \lambda_P e_P = \mu$$

which implies that λ_P is:

$$\lambda_P(e_R, e_P) = \frac{\mu - \varphi_R e_R}{e_P}$$

Given that the share of productive workers employed in the productive sector is $1 - \lambda_P$, total output produced in the economy is:

$$Y = \mu + (1 - \lambda_P(e_R, e_P))\, e_P = e_P + \varphi_R e_R$$

Suppose the government is utilitarian, the welfare function is:

$$W = e_P + \varphi_R e_R - h(e_P) - h(e_R)$$

If the government can observe and tax income through a non-linear tax schedule but cannot tax occupational choices, we can solve the problem by choosing an optimal effort level for each type. The FOCs for effort are:

$$h'(e_P) = 1$$

$$h'(e_R) = \varphi_R$$

Notice that the two conditions imply zero wedge on labor income for both types. Although rent-seekers are not productive at all they are not taxed in equilibrium. The reason is that in this model rent-seekers are "indirectly productive" by crowding out productive workers from the rent-seeking sector. If rent-seekers were taxed, productive workers would be attracted into the rent-seeking sector (e_R would fall but $E = \mu$, and λ_p would have to increase to balance the change) and would decrease total production. Notice that the result is not dependent on the assumption of utilitarian social preferences, but would hold for any other combination of social welfare weights.

This example shows how general equilibrium considerations might be extremely important in shaping optimal marginal tax rates. Even under the extreme assumption that all top earners are rent-seekers, general equilibrium considerations would put downward pressure on marginal tax rates at the top to avoid attracting productive workers into rent-seeking.

The model with occupational choices can also be employed to study the "trickle down" effects in Stiglitz (1982) (see Rotschild Scheuer, 2013). Allowing for occupational choices still pushes towards lower top marginal tax rate than in a standard Mirrlees model with linear production, but less so than in a world without occupational choice (as in Stiglitz 1982). The reason is that, unlike in the Stiglitz's model, if the government subsidizes high types, effort increases in the high skill sector decreasing wages in the high sector and increasing wages for the low sector. In a Roy model this would attract to the low-skilled sector some workers who were indifferent between the two sectors, reducing the increase in the low-skilled wage. This effect works against the standard general equilibrium effect presented in the previous paragraph, making trickle down less effective.

3 Wage Bargaining and Tax Avoidance

In this paragraph we study a standard Mirrlees model with a linear production function where individual income can depart from actual output. We present two potential departures: wage bargaining and tax avoidance.

Wage Bargaining Suppose top earners have measure 1 and after bargaining get a fraction η of their output z (where we allow for $\eta > 1$) such that $y = \eta z$. Bargained earnings are $b = (\eta - 1)\, y$ and average bargained earnings in the economy are $E\,(b)$. In the aggregate, it must be the case that total product is equal to total compensation. Hence, if $E\,(b) > 0$, so that there is overpay on average, $E\,(b)$ must come at the expense of somebody. The opposite is true if $E\,(b) < 0$. For simplicity, we assume that any gain made through bargaining comes uniformly at the expense of everybody else in the economy. Hence, individual incomes are all reduced by a uniform amount $E\,(b)$ if $E\,(b) > 0$.[2] We further assume that individuals can exert effort to increase η and their preferences are:

$$u_i\,(c, \eta, y) = c - h_i\,(y) - k_i\,(\eta)$$

When the cost of bargaining $E\,(b)$ is uniformly distributed across all agents, the government can offset it with the demogrant $-T\,(0)$. It follows that earnings can be written as $z = \eta y = y + b$. Each individual chooses η and z to maximize $u_i\,(c, \eta, y) = \eta y - T\,(\eta y) - h_i\,(y) - k_i\,(\eta)$ and FOCs are:

$$(1-\tau)\,\eta = h_i'\,(y)$$

$$(1-\tau)\,y = k_i'\,(\eta)$$

with $\tau = T'\,(z)$. Let us denote the average reported income, output and bargaining as Walrasian demands $z\,(1-\tau)$, $y\,(1-\tau)$ and $b\,(1-\tau)$. The implied elasticities are:

$$e_1 = \frac{1-\tau}{y}\frac{dy}{d\,(1-\tau)}$$

$$e = \frac{1-\tau}{z}\frac{dz}{d\,(1-\tau)}$$

$$e_2 = \frac{1-\tau}{z}\frac{db}{d\,(1-\tau)} = s\cdot e$$

with

$$s = \frac{db/d\,(1-\tau)}{dy/d\,(1-\tau)}$$

The definitions imply that $(y/z)\,e_1 = (1-s)\,e$ and that $e = (y/z)\,e_1 + e_3$.

When the social welfare weight on top incomes is zero, the government chooses the top tax rate to maximize total revenues:

$$\max_{\tau} \tau\,[z\,(1-\tau) - \bar{z}] - N\cdot E\,(b)$$

and the FOC is:

[2] This assumption can be relaxed allowing for income reductions affecting only workers in the same occupation using a framework similar to the one we presented in Section 6.

$$[z-\bar{z}]-\tau\frac{dz}{d(1-\tau)}=z-\bar{z}$$

rearranging

$$[\tau-s]\frac{dz}{d(1-\tau)}=z-\bar{z}\Rightarrow\frac{\tau-s}{1-\tau}\cdot e=\frac{z-\bar{z}}{z}=\frac{1}{a}$$

using $e_2=s\cdot e$ we get:

$$\tau e-e_2=\frac{1-\tau}{a}\Rightarrow\tau^*=\frac{1+e_2}{1+ae} \tag{7}$$

If top earners are overpaid relative to their productivity $s>0$ and $e_2>0$ implying that the optimal tax rate is higher than the one maximizing revenues in the standard model ($\tau^*>1/(1+ae)$). This is because of a *trickle up* effect that arise when a higher tax on high incomes reduces the cost of bargaining for low incomes. On the other hand, if $z<y$ and $e_2<0$ we would have a *trickle down* situation where a lower tax on top incomes shifts resources to individuals at the bottom.

Tax Avoidance Responses to tax rates can also take the form of tax avoidance. Define tax avoidance as changes in reported income due to changes in the form of compensation, but not in the total level of compensation. We observe tax-avoidance if taxpayers can shift part of their taxable income into another form that is treated more favorably from a tax perspective. Denote with x total sheltered income such that ordinary taxable income is $z=y-x$. Sheltered income is taxed at a constant marginal tax rate t. Suppose that the individual faces a utility cost for sheltering taxes and utility is $u_i(c,y,x)=c-h_i(y)-d_i(x)$ where $c=y-\tau z-tx+R=(1-\tau)y+(\tau-t)x+R$ and $R=\tau\bar{z}-T(\bar{z})$ is virtual income. We can write Walrasian demands $z(1-\tau,t)=y(1-\tau)-x(\tau-t)$. Let us define e_3 the elasticity of sheltered income:

$$e_3=\frac{1-\tau}{z}\frac{dx}{d(1-\tau)}=s\cdot e$$

where

$$s=\frac{dx/d(\tau-t)}{dy/d(1-\tau)+dx/d(\tau-t)}=\frac{dx/d(\tau-t)}{\partial z/\partial(1-\tau)}$$

and $e=(y/z)e_1+e_3$.

The government problem is:

$$\max_{\tau,t}\tau[z(1-\tau,t)-\bar{z}]+tx(\tau-t)$$

Suppose the government could only optimally set τ given some t, the FOC would be:

$$\begin{aligned}0&=[z-\bar{z}]-\tau\frac{\partial z}{\partial(1-\tau)}+t\frac{dx}{d(\tau-t)}\\&=[z-\bar{z}]-\tau\frac{\partial z}{\partial(1-\tau)}+ts\frac{\partial z}{\partial(1-\tau)}\\&=[z-\bar{z}]-ez\frac{\tau-ts}{1-\tau}\end{aligned}$$

rearranging:

$$\tau^*=\frac{1+t\cdot a\cdot e_2}{1+a\cdot e} \tag{8}$$

Notice how the tax is proportional to $t \cdot a \cdot e_2$ that captures the fiscal externality of tax avoidance. If $t = 0$ and the government cannot do anything to prevent income shifting, it is irrelevant whether e is due to real response or tax avoidance response (see Feldstein, 1999).

If instead the government could also optimally set t, we would have an extra optimality condition:

$$\begin{aligned} 0 &= \tau \frac{\partial z}{\partial t} + x - t \frac{dx}{d(\tau - t)} \\ &= x + (\tau - t) \frac{dx}{d(\tau - t)} \end{aligned}$$

since $x \geq 0$ and $dx/d(\tau - t) \geq 0$ the first order condition can only hold if $\tau = t$. If this is the case $x(\tau - t) = x(0) = 0$ and $z = y$ so that $e - e_3 = e_1$. If we replace this in (8) we obtain:

$$\tau^* = t^* = \frac{1}{1 + a \cdot e_1} \tag{9}$$

Intuitively, the government finds optimal to close any tax avoidance opportunity at the optimum. When this is the case the elasticity of income is the only one that matters.

References

[1] Feldstein, Martin. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." Review of Economics and Statistics 81 (4): 674–80.

[2] Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva. 2014. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." American Economic Journal: Economic Policy 6 (1): 230–71.

[3] Rothschild, Casey and Florian Scheuer, "Redistributive Taxation in the Roy Model," Quarterly Journal of Economics, 2013, 128, 623–668.

[4] Rothschild, Casey and Florian Scheuer, "Optimal Taxation with Rent-Seeking," Review of Economic Studies, 2016, forthcoming

[5] Scheuer, Florian (2014), Lecture Notes (find at `http://web.stanford.edu/~scheuer/top_incomes.pdf`)

[6] Stiglitz, Joseph, "Self-Selection and Pareto Efficient Taxation," Journal of Public Economics, 1982, 17, 213–240.