Topics in Dynamic Public Finance

John Hassler

Stockholm University

January 27, 2013

1 New Public Finance – the Mirrlees approach

The purpose of this part of the course is to study some simle public finance problems when there is heterogeneity in the population.These differences can be either in their productivity or in their value of leisure. Such differences imply that there is differences between individuals in their trade-off between leisure and work. It is assumed that the government cannot directly observe this differences, only observe the individuals market choices. For example, governments observe income, but not the effort exerted to get this income. The general problem is to redistribute and provide some public good. We will start by static examples and then go to some dynamic.

1.1 A standard static example Public Finance problem

We assume there is unit mass of individuals with different productivity, denoted θ. We normalize the average productivity to unity. Individuals derive utility from consumption, leisure and a public good. They have a standard utility function given by

$$U = u(c) - v(n) + \Gamma(G) \tag{1}$$

where c is consumption, n is labor supply (in units of effort or hours), G is a public good u and Γ are concave functions and v is convex.[1] We assume additive separability between private and public goods in order to abstract from differences between individual's taste for the public good coming from differences in their private consumption.

Often we will use the standard utility functions

$$u(c) = \frac{c^{1-\rho}}{1-\rho}$$
$$v(n) = \frac{n^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}$$

ρ is the constant relative risk aversion coefficient and γ the Frish labor supply elasticity (labor supply wage elasticity at constant marginal utility of consumption). Individuals are different in their productivities, denoted θ. We let $f(\theta)$ is the density of individual productivities. The aggregate resource constraint is then

$$\int_0^{\infty} f(\theta)(\theta n_\theta - c_\theta)\, d\theta - G = 0$$

where n_θ and c_θ represents effort and consumption by the type with productivity θ. The

[1] An alternative interpretation here is that $\Gamma(G)$ represents transfers to some non-working individuals.

planning problem is

$$\int_0^\infty f(\theta)\left(u(c_\theta) - \nu(n_\theta) + \Gamma(G)\right) + \lambda(\theta n_\theta - c_\theta - G)\, d\theta$$

FOC;

$$c_\theta; u'(c_\theta) = \lambda$$

$$n_\theta; v'(n_\theta) = \lambda\theta$$

$$\Gamma'(G) = \lambda$$

As we see, everyone consumes the same, but works depending on productivity. Since v is convex, $v'(n)$ is increasing and so is the inverse function $v^{-1\prime}(n_\theta)$ .Thus, higher produtivity individuals work more. With the utility specification above, we have

$$n_\theta = (\lambda\theta)^\gamma$$

With a Frisch elasticity of unity, labor supply is then proportional to productivity, not depending on the consumption elasticity. Why? An interesting implication of this is that a mean preserving spread in the distribution of productivities has no effect on consumption of private and public goods. With a lower elasticity, labor supply is concave in productivity implying that a mean preserving spread reduces labor supply.

Now, we have not yet looked at whether this allocation can be decentralized. Under which circumstances would it be possible?

The standard approach in PF is to endow the government with a set of arbitray tools to affect the market allocation and then let the government maximize over those. Such a problem is sometimes called a *Ramsey problem.*

Let's consider a very typical instrument. Namely a linear income tax τ. We then need to calculate the decentralized allocation as a function of τ. Now, each individual is choosing his labor supply to solve

$$\max\left(u\left(c_{\theta}\right)-v\left(n_{\theta}\right)+\Gamma\left(G\right)\right)$$

$$s.t. c_{\theta}=\left(1-\tau\right)\theta n_{\theta}$$

Note that we assume $\frac{\partial G}{\partial n_{\theta}}=0$, why?

Define the solution

$$n_{\theta}^{*}\left(\tau\right)\equiv\arg\max\left(u\left(\left(1-\tau\right)\theta n_{\theta}\right)-v\left(n_{\theta}\right)+\Gamma\left(G\right)\right)$$

and the indirect private utility

$$V_{\theta}\left(\tau\right)\equiv u\left(\left(1-\tau\right)\theta n_{\theta}^{*}\left(\tau\right)\right)-v\left(n_{\theta}^{*}\left(\tau\right)\right).$$

We can now set up the Ramsey problem as

$$L\equiv\int_{0}^{\infty}f\left(\theta\right)\left(V_{\theta}\left(\tau\right)+\Gamma\left(G\right)\right)d\theta$$

$$s.t. 0=\int_{0}^{\infty}f\left(\theta\right)\left(\theta n_{\theta}\left(\tau\right)\tau\right)d\theta-G$$

The first order condition for τ and G are

$$\frac{\partial}{\partial \tau}\int_0^{\infty} f(\theta) V_\theta(\tau) d\theta = -\lambda \frac{\partial}{\partial \tau}\int_0^{\infty} \theta n_\theta(\tau) \tau d\theta$$

$$\Gamma'(G) = \lambda$$

where λ is the shadow constraint on the resource constraint. This yields,

$$\frac{-\frac{\partial}{\partial \tau}\int_0^{\infty} f(\theta) V_\theta(\tau) d\theta}{\frac{\partial}{\partial \tau}\int_0^{\infty} \theta n_\theta(\tau) \tau d\theta} = \Gamma'(G)$$

For later use we note that we can write this as

$$\frac{-\int_0^{\infty} f(\theta) \frac{\partial V_\theta(\tau)}{\partial \tau} d\theta}{\int_0^{\infty} f(\theta) \frac{\partial(\theta n_\theta(\tau)\tau)}{\partial \tau} d\theta} = \Gamma'(G)$$

Define aggregate private indirect utility

$$U(\tau) \equiv \int_0^{\infty} f(\theta) V_\theta(\tau) d\theta$$

and

$$R(\tau) \equiv \int_0^{\infty} f(\theta)(\theta h_\theta(\tau)\tau) d\theta$$

is government reveneue.

Defining

$$MU(\tau) \equiv -\frac{dU}{d\tau} / \frac{dR}{d\tau}$$

as the marginal aggregate disutility of tax revenues, the optimality condition can therefore be written as the following equation in τ;

$$MU(\tau) = \Gamma'(R(\tau)) \tag{2}$$

i.e., that the aggregate utility loss of an extra dollar of revenue should be equal to the marginal value of the public good.

By specifying the functions, we can go further and do a quantitative analysis.

Let's use the specifications above. The problem of each individual is

$$\max \frac{(\theta(1-\tau)n_\theta)^{1-\rho}}{1-\rho} - \frac{n_\theta^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}$$

yielding

$$n_\theta = (\theta(1-\tau))^{\frac{\gamma(1-\rho)}{1+\rho\gamma}}$$

Note that with $\rho = 1$, utility is logaritmic in consumption and labor supply is independent of taxes and productivity (compare to first best). With ρ smaller (larger) than 1, labour supply increases (decreases) in net wage. Explain!

Using the solution in the two felicity functions yields private indirect private utility is then

$$V_\theta(\tau) = \left(\frac{1+\gamma\rho}{(1-\rho)(\gamma+1)}\right)(\theta(1-\tau))^{\frac{(1-\rho)(\gamma+1)}{1+\gamma\rho}}$$

and

$$\frac{\partial V_\theta(\tau)}{\partial\tau} = -\theta^{\frac{(1-\rho)(\gamma+1)}{1+\gamma\rho}}(1-\tau)^{\frac{(1-\rho)(\gamma+1)}{1+\gamma\rho}-1}.$$

Result .

1. *The marginal individual cost of taxation is decreasing (increasing) in ability θ if $\rho >$ $(<)\,1$.*

2. *The average marginal individual cost of taxation is increasing in ability inequality if $\rho > 1$ or smaller than $\frac{\gamma}{2\gamma+1}$.*

   The first part is obvious, the second follows since $-\frac{dV_\theta(\tau)}{d\tau}$ is then convex.

From this follows that if $\rho > 1$,

a agents with ability below that of the representative agent has a higher marginal utility loss of taxes than the representative agent,

b the average marginal utility loss of taxation is higher than that of the median agent (making a representative agent analysis problematic).

We can also compute the marginal revenue from a tax increase for each individual;

$$
\begin{aligned}
\frac{\partial\left(\theta n_\theta\left(\tau\right)\tau\right)}{\partial\tau} &= \theta n_\theta\left(\tau\right) + \theta\tau\frac{\partial n_\theta\left(\tau\right)}{\partial\tau} \\
&= \theta^{\frac{\gamma+1}{1+\gamma\rho}}\left(1-\tau\right)^{-\frac{1+2\gamma\rho-\gamma}{1+\gamma\rho}}\left(1-\tau\frac{1+\gamma}{1+\gamma\rho}\right).
\end{aligned}
$$

We see this is zero for an interior value of τ if $\frac{1+\gamma}{1+\gamma\rho} > 1$. This requires $\rho < 1$. Thus, we a riskaversion larger or equal to 1, there is no *Laffer curve* maximum. What happens if we use the tax revenue not for government consumption but for transfers?

Now assume that θ is log normal with mean μ and variance σ^2, the expectation of θ^x is $e^{\mu x + x^2\frac{\sigma^2}{2}}$. Under the normalization that average productivity is unity, i.e., $\bar{\theta} \equiv \int_0^\infty f\left(\theta\right)\theta d\theta =$

1, we must set

$$\mu = -\frac{\sigma^2}{2}$$

Using this, it is immediate that

$$\begin{aligned}\frac{\partial U(\tau)}{\partial \tau} &= \int_0^{\infty} f(\theta) \frac{\partial V_\theta(\tau)}{\partial \tau} d\theta \qquad (3)\\ &= (1-\tau)^{\frac{\gamma-\rho(1+2\gamma)}{1+\gamma\rho}} e^{-(1-\rho)(1+\gamma)\frac{\rho+\gamma(2\rho-1)}{(1+\gamma\rho)^2}\frac{\sigma^2}{2}}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial R(\tau)}{\partial \tau} &= \int_0^{\infty} f(\theta) \frac{\partial(\theta h_\theta(\tau)\tau)}{\partial \tau} d\theta \qquad (4)\\ &= (1-\tau)^{-\frac{1+2\gamma\rho-\gamma}{1+\gamma\rho}} \left(1-\tau\frac{1+\gamma}{1+\gamma\rho}\right) e^{(1-\rho)\gamma\frac{(1+\gamma)}{(1+\gamma\rho)^2}\frac{\sigma^2}{2}} \qquad (5)\end{aligned}$$

Then,

$$MU(\tau) = \frac{\frac{-\partial U(\tau)}{\partial \tau}}{\frac{\partial R(\tau)}{\partial \tau}} = \frac{1}{\left(1-\tau\frac{1+\gamma}{1+\gamma\rho}\right)}(1-\tau)^{\frac{1-\rho}{1+\gamma\rho}} e^{-\rho\frac{(1-\rho)(1+\gamma)(1+2\gamma)}{(1+\gamma\rho)^2}\frac{\sigma^2}{2}}$$

The standard log-normal distribution has only one parameter (σ) determining income dispersion and there is a simple one-to-one mapping between σ and the Gini-coefficient (ϕ);

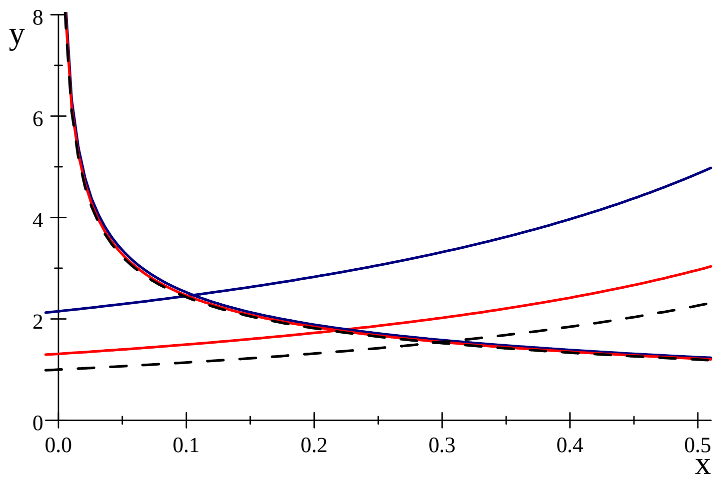
$$\phi = 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$$

where Φ is the standard normal cumulative distribution.[2] In the data, US has a Gini of 0.408, corresponding to $\sigma = 0.7579$. Sweden has a Gini of 0.25, corresponding to $\sigma = 0.4506$.[3] Now

[2] See Bourguignon (2003).

[3] Source: United Nations (2006). Table 15: Inequality in income or expenditure (PDF). Human Develop-

we can calculate the marginal cost of linear tax revenue. In the following graph, I have calculated the cost of revenue with $\rho = 2$ and $\gamma = \frac{1}{2}$. The utility of public good is set to $\Gamma(G) = k\frac{G^{1-\kappa}}{1-\kappa}$ and we use the budget constraint $G = R(\tau)$. We can calibrate k and κ by using spending shares on public goods $\tau_{US} = 10.54\%$, and $\tau_{SE} = 21.83\%$ reported in the Penn Word Table[4] for year 2004, giving $k = 0.97214$ and $\kappa = 0.40219$.



The downward sloping curves in Figure 1 are $\Gamma'(R(\tau))$ for the US and Swedish parametrizations. In fact, these curves are basically indistinguishable, indicating that government revenue as a function of the tax rate are very similar under the two different parameter sets. Instead, the large difference in chosen tax rates are due to differences in the utility loss associated with taxation. The upward-sloping solid curves are $MU(\tau)$ for the US and Sweden respectively. In the case of no heterogeneity, depicted by the dashed lines in Figure 2, the marginal utility cost of taxation is even lower than in Sweden. Therefore, taxes and public good provision should be higher. In fact, with no inequality, the optimal tax rate is

ment Report 2006 335. United Nations Development Programme.

[4]Source: Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

as high as 29.6%.

Finally, let's consider if the allocation is pareto efficient. Take a particular agent i in the productivity distribution. That person pays $\tau\theta_i h^*_{\theta_i}$ in taxes. Suppose now we ask this person to pay this amount as a lump sum payment but that then his marginal tax is zero. The problem is therefore

$$\max \left(u\left(c_{\theta_i}\right) - v\left(n_{\theta_i}\right) + \Gamma\left(G\right)\right)$$

$$s.t. c_\theta = \theta n_{\theta_i} - T$$

where T is exogeneously fixed to $\tau\theta_i n^*_{\theta_i}$. Clearly, the individual could still choose the "old" allocation, call it $c^*_{\theta_i}, h^*_{\theta_i}$. We know that

$$\frac{v'\left(n^*_{\theta_i}\right)}{u'\left(c^*_{\theta_i}\right)} = \theta_i\left(1-\tau\right)$$

Therefore, provided $\tau > 0$,

$$\frac{v'\left(n^*_{\theta_i}\right)}{u'\left(c^*_{\theta_i}\right)} < \theta_i.$$

So with a zero marginal tax and lump sum taxes, the individual would choose to work more (increasing v') and consume more (reducing u'). That would make him strictly better of and no one else worse off.

The problem, however, is that other people might want to have this deal to. In particular those with higher productivity. The ones with sufficiently low productivity would not like to pay $\tau\theta_i n^*_{\theta_i}$. This suggest a scheme where individuals are given a menue where lower marginal tax rates and higher transfers and then volontarily sort them selves. This is what the Mirrlees allocation achieves!

1.2 The static Mirrlees model

Consider now a simple two type variant of the model above. Furthermore disregard public good provision. Suppose a share π of the population has high productivity θ_h and the remaining share has productivity $\theta_l \leq \theta_h$. Consider first the first best allocation if the social welfare function is utilitarian

$$\max \pi \left(u\left(c_h\right) + v\left(n_h\right)\right) + (1-\pi)\left(u\left(c_l\right) + v\left(n_l\right)\right) \tag{6}$$
$$s.t. 0 \leq \pi\theta_h n_h + (1-\pi)\theta_l n_l - \pi c_h - (1-\pi) c_l$$

where subscripts denote the type, so c_h, for example, denoted consumption of the high productivity types.

Denoting the shadow value on the resource constraint by λ, we have the first order conditions

$$\pi u'(c_h) - \lambda\pi = 0$$
$$(1-\pi) u'(c_l) - (1-\pi)\lambda = 0$$
$$\pi v'(n_h) + \lambda\pi\theta_h = 0$$
$$(1-\pi) v'(n_l) + (1-\pi)\lambda\theta_l = 0$$
$$\lambda\left(\pi\theta_h n_h + (1-\pi)\theta_l n_l - \pi c_h - (1-\pi) c_l\right) \geq 0$$

Clearly the two first constraints imply that

$$c_h = c_l$$

while the next two implies

$$\frac{v'(n_h)}{v'(n_l)} = \frac{\theta_h}{\theta_l} \geq 1$$

that is the marginal disutility of work is higher for the able individuals, i.e., they work more. Clearly this poses a problem if the planner cannot observe individual productivity and the effort h the individual puts in. *The planner is assumed to only observe income and consumption.*

Furthermore,

$$\theta_h = \frac{-v'(n_h)}{u'(c_h)}$$
$$\theta_l = \frac{-v'(n_l)}{u'(c_l)}$$

with a well-known interpretation.

Consider now the problem of maximizing the utilitarian welfare function subject to the resource constraints and the incentive constraints, i.e., that individuals themselves choose labor supply and savings. A way of finding the second best allocation is to let the planner provide consumption and tell the individual to provide a given amount of income conditional on the ability an individual claims to have. So let's consider a situation where each individual reports her type and the planner then tells her how much income to provide y_i and how much to consume c_i. Let's call the report i_r. The incentive constraint is then that individuals voluntarily report their true ability. According to the *revelation principle,* any incentive compatible allocation can be achived in this way. Thus we can restrict ourselves to look within the class of allocations that satisfy incentive constraints. Later, we will discuss how to decentralize that, i.e., construct a tax-transfer system such that the optimal incentive-compatible allocation is chosen by the individuals.

The problem is now to solve (6) subject to the truth-telling constraint

$$u\left(c_i\right)+v\left(\frac{y_i}{\theta_i}\right) \geq u\left(c_{i_r}\right)+v\left(\frac{y_{i_r}}{\theta_i}\right), \forall i_r, i \in \{h, l\}$$

where we have substituted for n by y/θ. Note that we always divide by the true ability. Why?

We will not have both truth-telling constraints binding in the optimal allocation. We conjecture that truth-telling for the more able person binds. Why? Let's call the shadow value on that constraint by λ_I and the resource constraint λ_r. The problem is then

$$\max \pi\left(u\left(c_h\right)+v\left(\frac{y_h}{\theta_h}\right)\right)+(1-\pi)\left(u\left(c_l\right)+v\left(\frac{y_l}{\theta_l}\right)\right) \tag{7}$$

$$s.t. 0 \leq \pi y_h+(1-\pi) y_l-\pi c_h-(1-\pi) c_l$$

$$0=u\left(c_h\right)+v\left(\frac{y_h}{\theta_h}\right)-u\left(c_l\right)-v\left(\frac{y_l}{\theta_h}\right) \tag{8}$$

First order conditions are

$$\pi u'(c_h) - \lambda_r \pi + \lambda_I u'(c_h) = 0$$

$$(1-\pi) u'(c_l) - \lambda_r (1-\pi) - \lambda_I u'(c_l) = 0$$

$$\pi v'\left(\frac{y_h}{\theta_h}\right)\frac{1}{\theta_h} + \pi\lambda_r + \lambda_I v'\left(\frac{y_h}{\theta_h}\right)\frac{1}{\theta_h} = 0$$

$$(1-\pi) v'\left(\frac{y_l}{\theta_l}\right)\frac{1}{\theta_l} + (1-\pi)\lambda_r - \lambda_I v'\left(\frac{y_l}{\theta_h}\right)\frac{1}{\theta_h} = 0$$

These implies

$$\frac{u'(c_h)}{u'(c_l)} = \frac{1-\frac{\lambda_I}{1-\pi}}{1+\frac{\lambda_I}{\pi}}$$

Thus, the higher is the λ_I, the larger is the spread in marginal utilities.

Note also that

$$u'(c_h)\left(1+\frac{\lambda_I}{\pi}\right) = -v'\left(\frac{y_h}{\theta_h}\right)\frac{1}{\theta_h}\left(1+\frac{\lambda_I}{\pi}\right)$$

implying

$$\theta_h = \frac{-v'(n_h)}{u'(c_h)}$$

while

$$-\frac{v'\left(\frac{y_l}{\theta_l}\right)}{u'(c_l)} = \frac{1-\frac{\lambda_I}{1-\pi}}{1-\frac{\lambda_I}{(1-\pi)}\frac{v'\left(\frac{y_l}{\theta_h}\right)}{v'\left(\frac{y_l}{\theta_l}\right)}\frac{\theta_l}{\theta_h}}\theta_l < \theta_l$$

since $1 > \frac{v'\left(\frac{y_l}{\theta_h}\right)}{v'\left(\frac{y_l}{\theta_l}\right)}\frac{\theta_l}{\theta_h}$. Thus the labor leisure choice is distorted for the low ability types but not for the high ability types. The no distortion at the top is a quite general result when the distribution of abilities is bounded.

Take a simple example where $u(c) = \ln c$ and $v(n) = -\frac{n^2}{2}$. Set $\pi = 1/2$ and $\theta_h = 2, \theta_l = 1$.

Then, we have

$$
\begin{aligned}
\frac{1}{2}c_h^{-1} - \lambda_r \frac{1}{2} + \lambda_I c_h^{-1} &= 0 \\
\frac{1}{2}c_l^{-1} - \lambda_r \frac{1}{2} - \lambda_I c_l^{-1} &= 0 \\
-\frac{1}{2}n_h \frac{1}{2} + \frac{1}{2}\lambda_r - \lambda_I n_h \frac{1}{2} &= 0 \\
-\frac{1}{2}n_l + \frac{1}{2}\lambda_r + \lambda_I n_l \frac{1}{4} &= 0 \\
2n_h + n_l - c_h - c_l &= 0 \\
\ln c_h - \frac{n_h^2}{2} - \left(\ln (c_l) - \frac{\left(\frac{n_l}{2}\right)^2}{2} \right) &= 0
\end{aligned}
$$

The solution is: $n_l = 0.73338, \lambda_r = 0.68609, \lambda_I = 0.12896, c_h = 1.8334, c_l = 1.0816, n_h = 1.0908$

Note that $c_h n_h = 2 = \theta_h$, while $c_l n_l < 1 = \theta_l$.

In first best, we instead have

$$
\begin{aligned}
\frac{1}{2}c_h^{-1} - \lambda_r \frac{1}{2} &= 0 \\
\frac{1}{2}c_l^{-1} - \lambda_r \frac{1}{2} &= 0 \\
-\frac{1}{2}n_h \frac{1}{2} + \frac{1}{2}\lambda_r &= 0 \\
-\frac{1}{2}n_l + \frac{1}{2}\lambda_r &= 0 \\
2n_h + n_l - c_h - c_l &= 0
\end{aligned}
$$

with the solution is: $\{\lambda_r = 0.632\,46, c_h = 1.\,581\,1, c_l = 1.\,581\,1, n_h = 1.\,264\,9, n_l = 0.632\,46\}$, in which case $c_h n_h = \theta_h$ and $c_l n_l = \theta_l$.

1.2.1 Implementation

In the simple case discussed above, we can implement the allocation with a menue of marginal tax rates and transfers. Since the labor-leisure tradeoff is distorted (not distorted) for the low (high) ability individuals, we need a tax on labor for only the low ability type. For the low ability type to accept this, we need to give him a larger lump-sum transfer. Thus, indivduals are asked to choose either a positive marginal tax and a high transfer or a zero marginal tax and a smaller transfer (typically negative). Think of the intution for why this is optimal.

Given that the truth telling constraint is satisfied, individuals solve

$$\max \left(u\left(c_i\right)+v\left(n_i\right)\right)$$

$$s.t. c_i=\theta_i n_i\left(1-\tau_i\right)+T_i$$

Implying

$$\theta_i\left(1-\tau_i\right)=\frac{-v'\left(n_i\right)}{u'\left(c_i\right)}$$

In the example, we then have the two private first-order conditions and two budget constraints.

Plugging in the numbers and solving yields

$$[c_h n_h = \theta_h (1 - \tau_h)]_{n_h=1.0908, c_l=1.081\,6, n_l=0.733\,38, c_h=1.8334, \theta_h=2, \theta_l=1}$$

$$[c_l n_l = \theta_l (1 - \tau_l)]_{n_h=1.0908, c_l=1.081\,6, n_l=0.733\,38, c_h=1.8334, \theta_h=2, \theta_l=1}$$

$$[c_h = \theta_h n_h (1 - \tau_h) + T_h]_{n_h=1.0908, c_l=1.081\,6, n_l=0.733\,38, c_h=1.8334, \theta_h=2, \theta_l=1}$$

$$[c_l = \theta_l n_l (1 - \tau_l) + T_l]_{n_h=1.0908, c_l=1.081\,6, n_l=0.733\,38, c_h=1.8334, \theta_h=2, \theta_l=1}$$

The solution is: $\{T_h = -0.348\,06, T_l = 0.499\,87, \tau_l = 0.206\,78, \tau_h = 0\}$

Finally, we need to check whether it is necessary to add some non-linearities in the tex system. Consider the utility if the high transfer, high marginal tax is chosen by the high ability type. The choice then satisfies

$$[c_h n_h = \theta_h (1 - \tau_l)]_{\theta_h=2, \tau_l=0.20678, T_h=-0.34806, T_l=0.49987}$$

$$[c_h = \theta_h n_h (1 - \tau_l) + T_l]_{\theta_h=2, \tau_l=0.20678, T_h=-0.34806, T_l=0.49987}$$

with the solution $\{c_{hdev} = 1.8559, n_{hdev} = 0.85479\}$. Clearly, this gives higher utility and we need to prevent this deviation. This can be done by having another bracket in the tax system. The following tax system could then implement the optimal second-best allocation. The indivuals choose from the following menue;

1. A lump sum tax $-T_h = 0.348$. No marginal income tax.

2. A lum sum transfer $T_l = 0.500$. A marginal income tax of $\tau_l = 20.7\%$ up to income $n_l = 0.733$. Above that, a sufficently high tax rate to deter any benefit claimant to

earn more, e.g., 100%.

1.3 Uniform commodity taxation

An important assumption in the previous subsection was that there is just one good. In reality, there are many goods, both intermediaries and final goods. Then, a key issue becomes; Should different goods be taxed at different rates, i.e., should we use differentiated VAT's? If not, we have seen that it does not matter whether we use a flat consumption tax or a proportional income tax.

One of the most celebrated results in public finance is the Atkinson-Stiglitz *uniform commodity taxation result* (Atkinson & Stiglitz, 1972). This states that under some conditions, most importantly that utility is separable in leisure and an aggregate of market consumption goods, a uniform tax rate should be used. Then, it can, as we have discussed above be replaced by a uniform tax rate on labor income. Loosely speaking, separability means that utility can be written as a function of a consumption aggregate $g(c)$,where c is a vector $[c_1, ..., c_n]$ of consumption goods bought in the market, and labor n (equivalently, leisure) Thus

$$\bar{u}\left(c_1, ...c_n, l\right) = u\left(g\left(c\right), n\right).$$

As above, productivity is unobserved by the planner and he only observes total income, not wages. Due to separability, we can separate the consumers problem in two steps. The last is to maximize $g\left(c\right)$ over the different consumption goods, given disposable income ω and the prices q_i (including taxes).

$$\max_{c} g\left(c_1, ..., c_n\right) \tag{9}$$

$$s.t. \ \sum_{i} q_i c_i \leq \omega$$

This generates demand functions $d_i(q,\omega)$ and an associated value function $h\left(q,\omega\right) \equiv g\left(d(q,\omega)\right)$. The latter function h, can be thought of has the optimal consumption aggregate, given prices and income.

The first step is then to choose labor supply by solving

$$\max_{y} u\left(h\left(q,\omega(y)\right), \frac{y}{\theta}\right),$$

where $\omega\left(y\right)$ i disposable income given gross income y.

Let's follow Boadway and Pestieau (2002) and consider the case were there are two types, $i \in \{h,l\}$ with different planner unobserved productivities (wages), $\theta_h > \theta_l$. We assume that there are two consumption goods, c_1 and c_2 and normalize their relative market price before taxes to unity. Without loss of generality, we assume the policy instrument in terms of consumption taxes is the tax on good 2 and set the other consumption tax to zero. This is w.o.l.g. since a common tax is equivalent to a labor income tax. The price on good 2 faced by consumers is $1+\tau \equiv q$ implying that the budget constraint of the agent of type i is

$$\omega_i = c_1 + qc_2.$$

The second step problem (9) can now be written

$$h(q,\omega) = \max_{c_2} g(\omega - qc_2, c_2) \tag{10}$$

giving

$$\frac{g_2}{g_1} = q$$

and using the envelope theorem, we have

$$h_\omega = g_1, \tag{11}$$

$$h_q = -g_1 c_2 = -h_w c_2$$

We can now write the planner Lagrangian

$$L = \sum_{i=h,l} \pi_i u\left(h(q,\omega_i), \frac{y_i}{\theta_i}\right) + \lambda_r \sum_{i=h,l} \pi_i \left(y_i + \tau c_2^i - \omega_i\right)$$
$$+ \lambda_I \left(u\left(h(q,\omega_h), \frac{y_h}{\theta_h}\right) - u\left(\left(h(q,\omega_l), \frac{y_l}{\theta_h}\right)\right)\right)$$

The first constraint is the budget constraint of the government and the second is the incentive constraint. We conjecture as above that the high productivity type must be induced not to falsely report that he is a low productivity type.

Now, we focus on the FOC for the disposable incomes ω_i and the the consumer price q. To not have to write out the arguments of all functions, we use superscript on functions to denote type and hat's on functions denote for an h type who pretends to be of type l. We then get

$$\omega_l; \pi_l u_h^l h_\omega^l - \lambda_r \pi_l \left(1 - \tau \frac{\partial c_2^l}{\partial \omega_l}\right) - \lambda_I \hat{u}_h^h \hat{h}_\omega^h = 0$$
$$\omega_h; \pi_h u_h^h h_\omega^h - \lambda_r \pi_h \left(1 - \tau \frac{\partial c_2^h}{\partial \omega_h}\right) + \lambda_I u_h^h h_\omega^h = 0$$
$$q; \sum_{i=h,l} \pi_i u_h^i h_q^i + \lambda_r \sum_{i=h,l} \pi_i \left(c_2^i + \tau \frac{\partial c_2^i}{\partial q}\right) + \lambda_I \left(u_h^h h_q^h - \hat{u}_h^h \hat{h}_q^h\right) = 0$$

Now, multiply the first equation by c_2^l and the second by c_2^h and use (11). Giving

$$\omega_l; -\pi_l u_h^l h_q^l - \lambda_r \pi_l \left(1 - \tau \frac{\partial c_2^l}{\partial \omega_l}\right) c_2^l - \lambda_I \hat{u}_h^h \hat{h}_\omega^h c_2^l = 0$$
$$\omega_h; -\pi_h u_h^h h_q^h - \lambda_r \pi_h \left(1 - \tau \frac{\partial c_2^h}{\partial \omega_h}\right) c_2^h - \lambda_I u_h^h h_q^h = 0$$

Add these two to the FOC for q; This gives

$$\lambda_r \tau \sum_{i=h,l} \pi_i \left(\frac{\partial c_2^i}{\partial q} + \frac{\partial c_2^i}{\partial \omega_i} c_2^i\right) - \lambda_I \hat{u}_h^h \left(\hat{h}_\omega^h c_2^l + \hat{h}_q^h\right) = 0.$$

Now, consider the parenthesis in the second term, $\hat{h}_\omega^h c_2^l + \hat{h}_q^h$. Spelling out the arguments, we write this

$$h_\omega \left(q, \omega_l\right) c_2^l + h_q \left(\left(q, \omega_l\right)\right).$$

From (11) we know this is zero. Recall that this term comes from the cheating high productivity types, but since he consumes as much of good 2 as the the low productivity types, the same envelope condition holds. This would not be the case if also leisure entered in this expression, since the two types consume different amounts of leisure. We thus end up with

$$\lambda_r \tau \sum_{i=h,l} \pi_i \left(\frac{\partial c_2^i}{\partial q} + \frac{\partial c_2^i}{\partial \omega_i} c_2^i \right) = 0$$

Note that $\frac{\partial c_2^i}{\partial q} + \frac{\partial c_2^i}{\partial \omega_i} c_2^i$ is the derivative of the compensated demand function for c_2, i.e., the effect on demand of a marginal increase in the price dq together with an income transfer of dqc_2. Provided this is not zero, the tax must be zero.

The intuition for the result is that the planner wants to distort only margins that can help him identify the low productivity individuals (equivalently, the cheaters). If the marginal rate of substitution is the same for low and high productivity individuals for some pair of goods, there is no point in distorting it. One can, of course think of cases where this is not the case. For example, a cheating high productivity individual consumes a lot of leisure. Suppose there is one good that is a complement to leisure, like vacation trips. Such a good should then be taxed higher because it reduces the value of cheating for the high productivity individual.

A related result to the A-S is the Diamond-Mirrlees production efficiency result (Diamond & Mirrlees, 1972). This result states that production, in the sense the use of different inputs in production, should not be distorted. This result builds on a similar separability. If consumers care of the final product, not of how it is produced, distorting production cannot help the planner doing anything good.

1.4 The direct approach

An alternative to the Mirrleesian approach is to work directly with the tax system and derive optimal properties of that. Saez (2001) show that this can be done using observed characteristics as labor supply elasticites and the actual income distribution. To understand the intuition behind the fairly complicated formulas, consider a tax system $T(y)$ where y is gross income and $T(y)$ is the tax payment. Define $\tau(y) \equiv T'(y)$ and let $H(z)$ be the share of individuals with income at or below z, with a density denoted $h(z)$.

Consider the effects of a small increase in the marginal tax rate $d\tau$ over the small intervall y^* to $y^* + dy^*$. This change is illustrated in the figure below. Clearly, individuals with income below y^* are not affected by the change. Individuals in the interval $[y^*, y^* + dy^*]$ face a change in their marginal tax τ, but the average tax is (almost) not changed. Thus, there is only a substitution effect and the change in labor supply depends on the *compensated* income elasticity. Thus, an increase in that tax rate reduces labor supply. This is a negative effect seen from the point of view of a benevolent planner and the importance of it depends on the density of indivduals $h(y^*)$.

Above $y^* + dy^*$, the marginal income tax rate τ is unchanged but the average income is increased by $dy^* d\tau$. This has a mechanic direct effect on reveneues when behavioral changes are disregarded and an endogenous effect via labor supply that depends on the income elasticity of labor supply.

Assuming leisure is a normal good, the higher tax increases labor supply above y^*. Provided the value of government revenue is higher than the value of private spending for inviduals with income above y^*,both these effects are positive for the planner. The strength of them depends positively (loosely speaking) on total income above $y^* + dy^*$ and therefore on $(1 - H(y^* + dy^*)$.

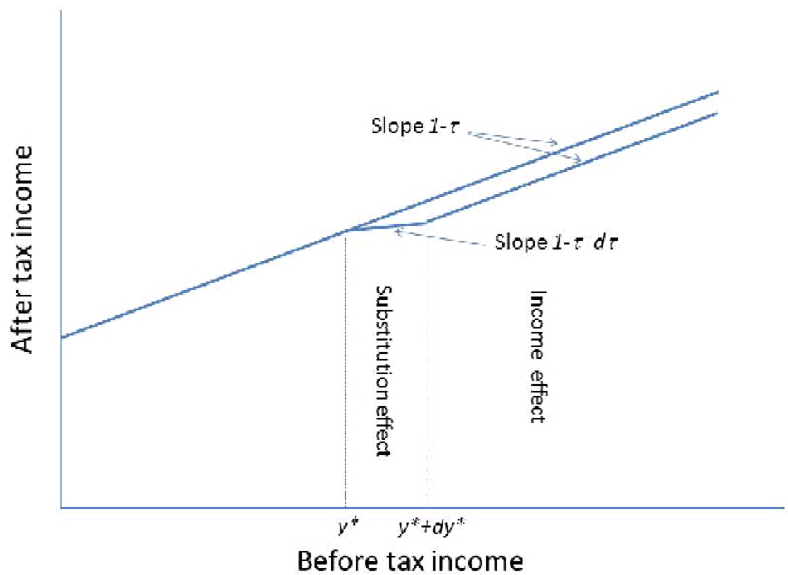
We have now defined positive and a negative effects of increasing the slope at y^*. If $T(y)$

is optimal, these effects should balance each other exactly. Furthermore, this should be true at all income levels y. Letting $dy^* \to 0$, this then defines a differential equation that must be satisfied. Together with, e.g., a financing reqirement or any other condition that pins down the total tax requirements, this defines the optimal tax. We note that the marginal tax $\tau(y)$, tends to be high;

- if the compensated elasticity at y is low,
- if $h(y)$ is low,
- if total income above y, i.e., $\int_z^\infty yh(y)\, dz$, is high.
- if income elasticity above y is high.
- if the planner's value of money is high relative to the value the planner attach to marginal income of individuals with income above y.

If there is a maximum income, the marginal tax rate should be zero there since no revenue is generated above this income. This is the zero tax rate at the top result of Mirrlees. In practice, however, we can not easily construct a system with a specific tax rate at the exact top.

Saez (2001) argues that empirically, shape of the income distribution affects the optimal tax rate at some each income level y by the term $\frac{1-H(y)}{yh(y)}$. The elastic distortion at y depends on the amount of income generated by individuals with income y, i.e., $yh(y)$. If this is large, distortions are large. Non-distortive revenues depends on how many individuals earn more than y, i.e., $1 - H(Y)$. Saez shows that the empirical income distribution is close to a Pareto



for high income, in which case $\frac{1-H(y)}{yh(y)}$ is constant. For low incomes, the value is high and for intermediate it is the lowest. In 1993, the minimum was at \$80 000. The optimal marginal tax rate is the initially high, decreasing to \$80000 than increasing to around \$200 000 and then constant. Of course, if we were to find the maximum, $1 - H(y) = 0$ and so the tax rate.

2 New Public Finance – the dynamic Mirrlees approach

Let us now consider the dynamic Mirrlees approach to optimal taxation. As above, individuals are assumed to be different. The only difference is that we know consider a dynamic environment.

Consider a simple two-period example.

Individual preferences are:

$$E\left(u\left(c_1\right)+v\left(n_1\right)+\beta\left(u\left(c_2\right)+v\left(n_2\right)\right)\right)$$

where c_t is consumption and n_t is labor supply/work effort. u is increasing and concave and v decreasing and concave. Individuals differ in their ability, denoted θ. It is assumed that there is a finite number $i \in \{1, 2, ..., N\}$ of ability levels and ability might change over time. We will interchangeably use type and ability to denote θ. Output is produced in competitive firms using a linear technology where each individual i produces

$$y_t\left(i\right)=\theta\left(i\right)n_t\left(i\right).$$

There is a large number of individuals of a unitary total mass. In the first period, individuals are given abilities by nature according to a probability function $\pi_1\left(i\right)$. The ability can then change to the second period. Second period ability is denoted $\theta\left(i,j\right)$ and the transition probability is $\pi_2\left(j|i\right)$.

There is a storage technology with return R. Finally, the government needs to finance some spendings G_1 and G_2. At first, we analyze the case of no aggregate uncertainty.

The aggregate resource constraint is

$$\sum_i\left(y_1\left(i\right)-c_1\left(i\right)+\sum_j\frac{y_2\left(i,j\right)-c_2\left(i,j\right)}{R}\pi_2\left(j|i\right)\right)\pi_1\left(i\right)+K_1=G_1+\frac{G_2}{R}\tag{12}$$

where K_1 is an aggregate initial endowment.

The problem is now to maximize the utilitarian welfare function subject to the resource constraints and the incentive constraints, i.e., that individuals themselves choose labor supply and savings. A way of finding the second best allocation is to let the planner provide consumption and work conditional on the ability an individual claims to have (and if relevant, the aggregate state). Here this is in the first period $c_1(i), y_1(i)$ and in the second, $c_2(i,j), y_1(i,j)$. Individuals then report their abilities to the planner. The strategy of an individual is his first period report and then a reporting plan as a function of the realized period 2 ability. Let's call the report i_r and $j_r(j)$, where the latter is the report as a function of the true ability. The incentive constraint is then that individuals voluntarily report their true ability. According to the *revelation principle,* this always yields the best incentive compatible allocation. The *truth-telling* constraint is then that

$$
\begin{aligned}
& u(c_1(i)) + v\left(\frac{y_1(i)}{\theta_1(i)}\right) + \beta \sum_j \left(u(c_2(i,j)) + v\left(\frac{y_2(i,j)}{\theta_2(i,j)}\right)\right)\pi_2(j|i) \qquad (13)\\
& \geq u(c_1(i_r)) + v\left(\frac{y_1(i_r)}{\theta_1(i)}\right) + \beta \sum_j \left(u(c_2(i_r, j_r(j))) + v\left(\frac{y_2(i_r, j_r(j))}{\theta_2(i,j)}\right)\right)\pi_2(j|i)
\end{aligned}
$$

for any possible reporting strategy $i_r, j_r(j)$. Note that the θ_s are the true ones in both sides of the inequality. Note also that *truth-telling* implies that

$$
u(c_2(i,j)) + v\left(\frac{y_2(i,j)}{\theta_2(i,j)}\right) \geq u(c_2(i_r, j_r(j))) + v\left(\frac{y_2(i_r, j_r(j))}{\theta_2(i,j)}\right) \forall j, \qquad (14)
$$

otherwise utility could be increased by reporting j_r if the second period ability is j. The planning problem is to maximize

$$
\sum_i \left(u(c_1(i)) + v\left(\frac{y_1(i)}{\theta_1(i)}\right) + \beta \sum_j \left(u(c_2(i,j)) + v\left(\frac{y_2(i,j)}{\theta_2(i,j)}\right)\right)\pi_2(j|i)\right)\pi(i)
$$

subject to (12) and (13).

Letting stars * denote optimal allocations. We can now define three wedges (distortions) that the informational friction may cause. These are the consumption-leisure (intratemporal) wedges

$$
\begin{aligned}
\tau_{y_1}(i) &\equiv 1 + \frac{v'\left(\frac{y_1^*(i)}{\theta_1(i)}\right)}{\theta_1(i)\, u'\left(c_1^*(i)\right)}, \\
\tau_{y_2}(i,j) &\equiv 1 + \frac{v'\left(\frac{y_2^*(i,j)}{\theta_2(i,j)}\right)}{\theta_2(i,j)\, u'\left(c_2^*(i,j)\right)},
\end{aligned}
$$

and the intertemporal wedge

$$
\tau_k(i) \equiv 1 - \frac{u'\left(c_1^*(i)\right)}{\sum_j \beta R u'\left(c_2(i,j)\right) \pi_2(j|i)}.
$$

Clearly, in absence of government interventions, these wedges would be zero by perfect competition and the first-order conditions of private optimization.

2.1 The inverse Euler equation

We will now show that if individual productivities are not always constant over time, the intertemporal wedge will not be zero. The logic is as follows and similar to what we have done above. In an optimal allocation, the resource cost (expected present value of consumption) of providing the equilibrium utility to each type, must be minimized. Consider the following perturbation around the optimal allocation for a given first period ability type i. Increase utility by a marginal amount Δ for all possible second period types $\{i,j\}$ the agent could become. To compensate, decrease utility by $\beta\Delta$ in the first period.

First, note that expected utility is not changed.

Second, since utility is changed in parallel for all ability levels the individual could have in the second period, their relative ranking cannot change. In other words, if we add Δ to both sides of (14) it must still be satisfied.

Thus, the incentive constraint is unchanged. However, the resource constraint is not necessarily invariant to this peturbation. Let

$$\begin{aligned}\tilde{c}_1(i;\Delta) &= u^{-1}\left(u\left(c_1^*(i)\right)-\beta\Delta\right),\\ \tilde{c}_2(i,j;\Delta) &= u^{-1}\left(u\left(c_2^*(i,j)\right)+\Delta\right)\end{aligned}$$

denote the perturbed consumption levels. The resource expected resource cost of these are

$$\begin{aligned}&\tilde{c}_1(i;\Delta)+\sum_j \frac{1}{R}\tilde{c}_2(i,j;\Delta)\,\pi_2(j|i)\\ &= u^{-1}\left(u\left(c_1^*(i)\right)-\beta\Delta\right)+\sum_j \frac{1}{R}u^{-1}\left(u\left(c_2^*(i,j)\right)+\Delta\right)\pi_2(j|i).\end{aligned}$$

The first-order condition for minimizing the resource cost over Δ must be satisfied at $\Delta = 0$, for the * consumption levels to be optimal.

Thus,

$$
\begin{aligned}
0 &= \\
&= \frac{-\beta}{u'\left(c_1^*(i)\right)} + \sum_j \frac{1}{R} \frac{1}{u'\left(c_2^*(i,j)\right)} \pi_2(j|i) \\
&\Rightarrow \frac{1}{u'\left(c_1^*(i)\right)} = E_1 \frac{1}{\beta R u'\left(c_2^*(i,.)\right)},
\end{aligned}
$$

which we note is an example of the *inverse Euler equation.*

From Jensen's inequality, we find that

$$
\begin{aligned}
u'\left(c_1^*(i)\right) &< E\beta R u'\left(c_2^*(i,.)\right) \\
&\Rightarrow \tau_k(i) > 0,
\end{aligned}
$$

if and only if there is some uncertainty in c_2^*. Note that this uncertainty would come from second period ability being random and the allocation implying that second period consumption depends on the realization of ability. If second period ability is non-random, i.e., $\pi_2(j|i) = 1$ for some j, then $\tau_k(i) = 0$.

2.2 A simple logarithmic example: insurance against low ability.

Suppose in the first period, ability is unity and in the second $\theta > 1$ or $\frac{1}{\theta}$ with equal probability.Disregard government consumption – set $G_1 = G_2 = 0$, although non-zero spending is quite easily handled. The problem is therefore to provide a good insurance against a low-ability shock when this is not observed.

The first best allocation is the solution to

$$\max_{c_1, y_1, c_h, c_l, y_h, y_l} u\left(c_1\right) + v\left(y_1\right) + \beta\left(\frac{u\left(c_h\right) + v\left(\frac{y_h}{\theta}\right)}{2} + \frac{u\left(c_l\right) + v\left(\frac{y_l}{\frac{1}{\theta}}\right)}{2}\right)$$

$$s.t. 0 = y_1 + \frac{y_h + y_l}{2R} - c_1 - \frac{c_h + c_l}{2R}$$

First order conditions are

$$u'\left(c_1\right) = \lambda$$

$$v'\left(y_1\right) = -\lambda$$

$$\beta u'\left(c_h\right) = \frac{\lambda}{R}$$

$$\beta u'\left(c_l\right) = \frac{\lambda}{R}$$

$$\beta v'\left(\frac{y_h}{\theta}\right)\frac{1}{\theta} = -\frac{\lambda}{R}$$

$$\beta v'\left(\theta y_l\right)\theta = -\frac{\lambda}{R}$$

2.2.1 A simple example

Suppose for example that $u(c) = \ln(c)$ and $v(n) = -\frac{n^2}{2}$ and $\beta = R = 1$. Then, we get

$$\begin{aligned}
\frac{1}{c_1} &= \lambda \\
\frac{1}{c_h} &= \lambda \\
\frac{1}{c_l} &= \lambda \\
y_1 &= \lambda \\
\frac{y_h}{\theta^2} &= \lambda \\
y_l \theta^2 &= \lambda \\
c_1 + \frac{c_h + c_l}{2} - y_1 - \frac{y_h + y_l}{2} &= 0
\end{aligned}$$

We see immediately that $c_1 = c_h = c_l$ while $y_h = \theta^2 y_1$ and $y_l = \frac{y_1}{\theta^2}$ and $y_1 = \sqrt{\frac{2}{\left(1+\frac{1}{2}\left(\theta^2+\theta^{-2}\right)\right)}} = n_1$. Therefore, $n_h = \frac{y_h}{\theta} = \theta n_1$ and $n_l = y_l \theta = \frac{n_1}{\theta}$. Thus, if the individual becomes of high ability in the second period, he should work more but don't get any higher consumption. Is this incentive compatible?

We conjecture that the binding incentive constraint is for the high ability type. High has to be given sufficient consumption to make him voluntarily choose not to report being low ability. If he misreports, he gets c_l and is asked to produce y_l. The constraint is therefore

$$\begin{aligned}
& u(c_1) + v(y_1) + \beta \left(\frac{u(c_h) + v\left(\frac{y_h}{\theta}\right)}{2} + \frac{u(c_l) + v(\theta y_l)}{2} \right) \\
& \geq u(c_1) + v(y_1) + \beta \left(\frac{u(c_l) + v\left(\frac{y_l}{\theta}\right)}{2} + \frac{u(c_l) + v(\theta y_l)}{2} \right)
\end{aligned}$$

$$u\left(c_h\right)+v\left(\frac{y_h}{\theta}\right) \geq u\left(c_l\right)+v\left(\frac{y_l}{\theta}\right)$$
$$\ln c_h-\ln c_l \geq \frac{y_h^2-y_l^2}{2\theta^2}$$

We conjecture this is binding. The problem is then

$$\max_{c_1,y_1,c_h,c_l,y_h,y_l} \ln\left(c_1\right)-\frac{y_1^2}{2}+\left(\frac{\ln c_h-\frac{\left(\frac{y_h}{\theta}\right)^2}{2}}{2}+\frac{\ln c_l-\frac{\left(\theta y_l\right)^2}{2}}{2}\right)$$
$$s.t. 0=y_1+\frac{y_h+y_l}{2}-c_1-\frac{c_h+c_l}{2}$$
$$0=\ln c_h-\ln c_l-\frac{y_h^2-y_l^2}{2\theta^2}.$$

Denoting the shadow values by λ_r and λ_I the FOCs for the consumption levels are

$$c_1=\frac{1}{\lambda_r}$$
$$c_h=\frac{1+2\lambda_I}{\lambda_r}$$
$$c_l=\frac{1-2\lambda_I}{\lambda_r}$$

from which we see

$$\frac{c_h^*}{c_1^*}=1+2\lambda_I, \frac{c_l^*}{c_1^*}=1-2\lambda_I$$

and

$$\tau_k \equiv 1-\frac{u'\left(c_1^*\right)}{\beta R\left(\frac{u'\left(c_h^*\right)}{2}+\frac{u'\left(c_l^*\right)}{2}\right)}=1-\frac{\lambda_r}{\frac{\lambda_r}{1+2\lambda_I}\frac{1}{2}+\frac{\lambda_r}{1-2\lambda_I}\frac{1}{2}}=\left(2\lambda_I\right)^2,$$

implying a positive intertemporal wedge if the IC constraint binds.

The intratemporal wedges are found by analyzing the FOC's for the labor supplies

$$
\begin{aligned}
y_1^* &= \lambda_r \\
y_h^* &= \frac{\lambda_r}{1+2\lambda_I}\theta^2 \\
y_l^* &= \frac{\lambda_r}{\theta^4 - 2\lambda_I}\theta^2
\end{aligned}
$$

$$
\begin{aligned}
\tau_{y_1} &= 1 + \frac{v'\left(y_1^*\right)}{u'\left(c_1^*\right)} = 1 - \frac{y_1^*}{\frac{1}{c_1^*}} = 1 - \frac{\lambda_r}{\frac{1}{\frac{1}{\lambda_r}}} = 0, \\
\tau_{y_2}(h) &= 1 + \frac{v'\left(\frac{y_h^*}{\theta}\right)}{\theta u'\left(c_h^*\right)} = 1 + \frac{-\frac{y_h^*}{\theta}}{\theta\frac{1}{c_h^*}} \\
&= 1 + \frac{-\frac{\frac{\lambda_r}{1+2\lambda_I}\theta^2}{\theta}}{\theta\frac{1}{\frac{1+2\lambda_I}{\lambda_r}}} = 0
\end{aligned}
$$

and

$$
\begin{aligned}
\tau_{y_2}(l) &= 1 + \frac{v'\left(\theta y_l^*\right)}{\frac{1}{\theta}u'\left(c_l^*\right)} = 1 + \frac{-\theta y_l^*}{\frac{1}{\theta}\frac{1}{c_h^*}} \\
&= 1 + \frac{-\theta\frac{\lambda_r}{\theta^4-2\lambda_I}\theta^2}{\frac{1}{\theta}\frac{1}{\frac{1-2\lambda_I}{\lambda_r}}} = 2\lambda_I\frac{\theta^4-1}{\theta^4-2\lambda_I} > 0
\end{aligned}
$$

As we see, the wedge for the high ability types is zero, but positive for the low ability type.[5] For later use, we note that

$$\begin{aligned}
y_1^* c_1^* &= 1 \qquad (15)\\
y_h^* c_h^* &= \frac{\lambda_r}{1+2\lambda_I}\theta^2 \frac{1+2\lambda_I}{\lambda_r} = \theta^2\\
y_l^* c_l^* &= \frac{\lambda_r}{\theta^4 - 2\lambda_I}\theta^2 \frac{1-2\lambda_I}{\lambda_r} = \frac{1-2\lambda_I}{\theta^2\left(1-2\lambda_I\theta^{-4}\right)}
\end{aligned}$$

2.3 Implementation

It is tempting to interpret the wedges as taxes and subsidies. However, this is not entirely correct since the wedges in general are functions of all taxes. Furthermore, while there is typically a unique set of wedges this is generically not true for the taxes. As we have discussed above, many different tax systems might implement the optimal allocation. One example is the draconian, use 100% taxation for every choice except the optimal ones.

Only by putting additional restrictions is the implementing tax system found. Let us consider a combination if linear labor taxes and savings taxes that together with type specific transfers implement the allocation in the example. To do this, consider the individual

[5] The wedge, asymptotes to infinity as λ_I approach $\frac{\theta^4}{2}$. Can you explain?

problem,

$$\max_{c_1,y_1,s,y_h,y_l,c_h,c_l} \ln(c_1) - \frac{y_1^2}{2} + \left(\frac{\ln c_h - \frac{\left(\frac{y_h}{\theta}\right)^2}{2}}{2} + \frac{\ln c_l - \frac{(\theta y_l)^2}{2}}{2} \right)$$

$$s.t. 0 = y_1 (1 - \tau_1) - c_1 - s + T$$

$$0 = y_h (1 - \tau_h) + s (1 - \tau_{s,h}) - c_h + T_h$$

$$0 = y_l (1 - \tau_h) + s (1 - \tau_{s,l}) - c_l + T_l$$

with Lagrange multipliers λ_1, λ_h and λ_r.

First order conditions for the individuals are;

$$\begin{aligned}
\frac{1}{c_1} &= \lambda_1 \\
y_1 &= \lambda_1 (1 - \tau_1) \\
\lambda_1 &= \lambda_h (1 - \tau_{s,h}) + \lambda_l (1 - \tau_{l,h}) \\
\frac{y_h}{2\theta^2} &= \lambda_h (1 - \tau_h) \\
\frac{\theta^2 y_l}{2} &= \lambda_l (1 - \tau_l) \\
\frac{1}{2c_h} &= \lambda_h \\
\frac{1}{2c_l} &= \lambda_l
\end{aligned} \tag{16}$$

Using this, we see that

$$\frac{1}{c_1} = \frac{1}{2c_h} (1 - \tau_{s,h}) + \frac{1}{2c_l} (1 - \tau_{l,h})$$

Setting,

$$\tau_{s,h} = -2\lambda_I$$

$$\tau_{s,l} = 2\lambda_I.$$

this gives

$$\frac{1}{c_1} = \frac{1}{2c_h}\left(1+2\lambda_I\right) + \frac{1}{2c_l}\left(1-2\lambda_I\right)$$

which is satisfied if we plug in the optimal allocation $c_h^* = c_1^*\left(1+2\lambda_I\right)$ and $c_l^* = c_1^*\left(1-2\lambda_I\right)$

$$\frac{1}{c_1^*} = \frac{1+2\lambda_I}{2c_1^*\left(1+2\lambda_I\right)} + \frac{1-2\lambda_I}{2c_1^* 1-2\lambda_I}$$

Note that the expected capital income tax rate is zero, but it will make savings lower than without any taxes. Why?

Similarly, by noting from (15) that in the optimal second best allocation, we want

$$y_1 c_1 = y_1^* c_1^* = 1,$$

which is implemented by $\tau_1 = 0$. For the high ability type, the second best allocation in (15) is that $y_h^* c_h^* = \theta^2$, which is implemented by $\tau_h = 0$ since (18) implies that $y_h c_h = \theta^2\left(1-\tau_h\right)$.

For the low ability type, we want $y_l^* c_l^* = \frac{1-2\lambda_I}{\theta^2\left(1-2\lambda_I\theta^{-4}\right)}$. From (18), we know $y_l c_l = \frac{1-\tau_l}{\theta^2}$,so we solve

$$\frac{1-\tau_l}{\theta^2} = \frac{1-2\lambda_I}{\theta^2\left(1-2\lambda_I\theta^{-4}\right)}$$
$$\Rightarrow \tau_l = 2\lambda_I \frac{\theta^4-1}{\theta^4-2\lambda_I}.$$

Note that if $\lambda_I = \frac{1}{2}$, $\tau_l = 1$. I.e., the tax rate is 100%. There is no point going higher than that, so λ_I cannot be higher than $\frac{1}{2}$.

Finally, to find the complete allocation, we use the budget constraints of the private individual and the aggregate resource constraint. This will recover the transfers T, T_h and T_l. We should note that $T_l > T_h$ is consistent with incentive compatibility. Why? Because if you claim to be a low ability type you will have to may a high labor income tax which is bad if you are high ability and earn a high income. Thus, by taxing high income lower, we can have a transfer system that transfers more to the low ability types.

2.3.1 Third best – laissez faire.

The allocation in without any government involvements is easily found by setting all taxes to zero.

$$
\begin{aligned}
\frac{1}{c_1} &= \lambda_1 \\
y_1 &= \lambda_1 \\
\lambda_1 &= \lambda_h + \lambda_l \\
\frac{y_h}{2\theta^2} &= \lambda_h \\
\frac{\theta^2 y_l}{2} &= \lambda_l \\
\frac{1}{2c_h} &= \lambda_h \\
\frac{1}{2c_l} &= \lambda_l
\end{aligned} \tag{17}
$$

Using these and the budget constraints, we get

$$y_1 = \frac{1}{c_1}$$
$$\frac{1}{c_1} = \frac{1}{2c_h} + \frac{1}{2c_l}$$
$$\frac{y_h}{2\theta^2} = \frac{1}{2c_h}$$
$$\frac{\theta^2 y_l}{2} = \frac{1}{2c_l}$$

$$y_1 = c_1 + s$$

$$y_h + s = c_h$$

$$y_l + s = c_l$$

which implies

$$c_1 + s = \frac{1}{c_1}$$
$$\frac{1}{c_1} = \frac{1}{2c_h} + \frac{1}{2c_l}$$
$$c_h = \frac{1}{2}s + \frac{1}{2}\sqrt{s^2 + 4\theta^2}$$
$$c_l = \frac{\frac{1}{2}s\theta + \frac{1}{2}\sqrt{s^2\theta^2 + 4}}{\theta}$$

I did not find an analytical solution to this, but setting $\theta = 1.1$ I found the solution $c_1 = 0.997\,75, c_h = 1.102\,3, s = 4.\,504\,5 \times 10^{-3}, c_l = 0.911\,35,\ y_1 = 1.0023, y_h = 1.106\,8, y_l = 0.915\,85$.

As we see, consumption is lower in the first period and labor supply is higher than in second best. Consumption of high ability types is higher and labor supply lower than in second best. For low ability types, consumption is actually higher in *laissez faire* but also labor supply. The second period welfare of low ability types is higher in second best (-0.285 vs. $-0.300\,15$).

2.3.2 Means tested system

Suppose now we want to implement the optimal allocation without a savings-tax but using an asset tested disability transfer instead. That is we set

$$T_l = \begin{cases} T_l \text{ if } s \leq \bar{s} \\ -\bar{T} \text{ else.} \end{cases}$$

where $\bar{T}$ is sufficiently large to deter savings above $\bar{s}$. We set $\bar{s}$ equal to the first best $y_1^* - c_1^*$. Without a savings tax, the cap on savings will clearly bind due to the inverse Euler equation. The problem of the individual is therefore

$$\max_{c_1, y_1, s, y_h, y_l, c_h, c_l} \ln(c_1) - \frac{y_1^2}{2} + \left(\frac{\ln c_h - \frac{\left(\frac{y_h}{\theta}\right)^2}{2}}{2} + \frac{\ln c_l - \frac{(\theta y_l)^2}{2}}{2} \right)$$

$$s.t. 0 = y_1 (1 - \tau_1) - c_1 - \bar{s} + T$$

$$0 = y_h (1 - \tau_h) + \bar{s} - c_h + T_h$$

$$0 = y_l (1 - \tau_l) + \bar{s} - c_l + T_l$$

First order conditions for the individuals are;

$$\begin{aligned}
c_1; \frac{1}{c_1} &= \lambda_1 \\
y_1; y_1 &= \lambda_1 (1 - \tau_1) \\
y_h; \frac{y_h}{2\theta^2} &= \lambda_h (1 - \tau_h) \\
y_l; \frac{\theta^2 y_l}{2} &= \lambda_l (1 - \tau_l) \\
c_h; \frac{1}{2c_h} &= \lambda_h \\
c_l; \frac{1}{2c_l} &= \lambda_l
\end{aligned} \tag{18}$$

giving

$$1 - \tau_1 = c_1 y_1 \tag{19}$$

$$\theta^2 (1 - \tau_h) = c_h y_h \tag{20}$$

$$\frac{(1 - \tau_l)}{\theta^2} = c_l y_l$$

We want

$$1 = c_1 y_1 \Rightarrow \tau_1 = 0.$$

We also want

$$\begin{aligned}
c_h y_h &= \theta^2, \\
c_l y_l &= \frac{1 - 2\lambda_I}{\theta^2 (1 - 2\lambda_I \theta^{-4})}
\end{aligned} \tag{21}$$

requiring

$$\tau_h = 0,$$

$$\tau_l = 2\lambda_I \frac{\theta^4 - 1}{\theta^4 - 2\lambda_I},$$

mimicing the results above.

Golosow and Tsyvinski (2006), extend this model and calibrate it to the US. They assume people live until 75 years and start working at 25. The calibrate the probability of becoming permanently disabled for each age group. The problem is substantially simplified by the assumption that disability is permanent. They find the second best allocation in the same way as we have done here working backwards from the last period. As here, they show that the optimal allocation is implementable with transfers with asset limits and taxes on working people. The able should have zero marginal income taxes as in our example. In contrast to our example, the low ability types here have zero labor income and thus face no labor income tax.

An important finding is that asset limits are age dependent and increasing over (most of) the working life.

2.4 Time consistency

Under the Mirrlees approach, the government announces a menu of taxes or of consumption baskets. People then make choices that in equilibrium reveal their true types (abilities) to the government. Suppose the government could then re-optimize. Would it like to do this?

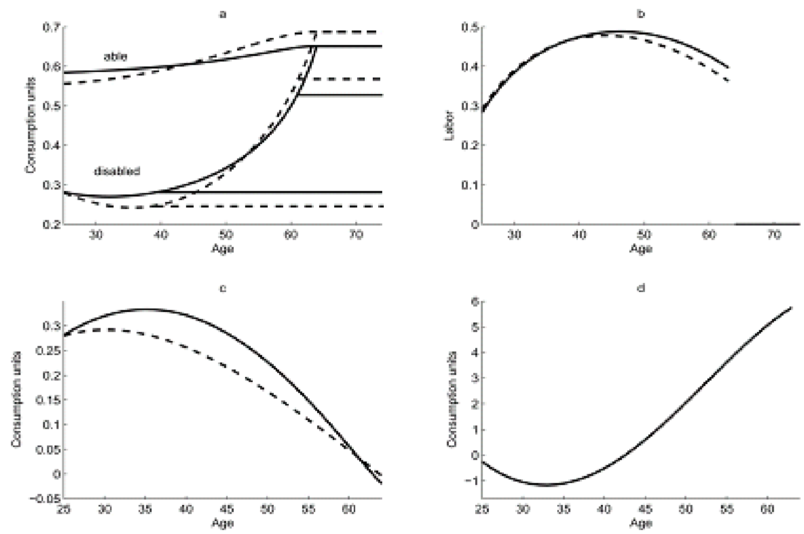


FIG. 2—Optimal disability programs with asset testing (solid lines) and without asset testing (dashed lines): *a*, consumption; *b*, labor; *c*, disability transfers; *d*, asset limits.

Figure 1: Figure from Golosov & Tsyvinski (2006)

The problem is more severe in a dynamic setting provided abilities are persistent. Why?

In a finite horizon economy, there might only be very bad equilibria (Roberts, 84). But better equilibria might arise in an infinite horizon setting.