Public Economics, Econ 8801
The Public Finance of Redistribution

Larry E. Jones, Keyvan Eslami, Anderson Schneider and Ali Shourideh*

1 A Simple, Static Model of Labor Supply

A continuum of mass one of individual agents, with identical preferences over consumption, c, and hours worked, l, given by $U(c,l)$ (increasing and concave, decreasing and convex; i.e. $V(c,\ell) \equiv U(c,1-\ell)$ is increasing and concave in (c,ℓ) where $\ell = 1-l$ is leisure). Each agent has a distinctive labor productivity, θ, so that, if l hours are worked, an agent of type θ produces θl units of effective labor. The distribution of θ's across agents is given by the probability $dG(\theta)$.[1]

1.1 Competitive Equilibrium Version: Ex-Post

There is a consumption good, and the different types of labor. We will use the consumption good, c, as numeraire.

The agent of productivity type θ solves:

$$\begin{aligned} \max_{c,l} \quad & U(c,l) \\ s.t. \quad & c \leq w(\theta)\, l. \end{aligned}$$

We will assume that there is a constant returns to scale production function in the input *vector* $l_f(\theta)$, where $l_f(\theta)$ is the number of hours worked by each worker of type θ. Then, the firm's problem is:

$$\begin{aligned} \max_{y,l_f(\theta)} \quad & y - \int w(\theta)\, l_f(\theta)\, dG(\theta) \\ s.t. \quad & y \leq \int \theta l_f(\theta)\, dG(\theta). \end{aligned}$$

*Keyvan has made numerous improvements throughtout this document. Ali provided the section on the Continuous Type Case and Anderson wrote the Appendix.

[1] Henceforth, we use l to denote the supply of labor, whereas ℓ is reserved for the consumption of leisure.

Hence, profit maximization requires that $w(\theta) = \theta$ for all θ.

Thus, an equilibrium can be described as follows.

Definition 1 *An ex post competitive equilibrium is given by a wage function, $w(\theta)$, and an all allocation for each type, $(c(\theta), l(\theta))$ (why must each agent of type θ get the same thing?), and an allocation for the representative firm, $(y, l_f(\theta))$, such that:*

1. *for each θ, $(c(\theta), l(\theta))$ solves:*

$$\begin{aligned} \max_{c,l} \quad & U(c,l) \\ s.t. \quad & c \leq w(\theta) l, \end{aligned}$$

2. *$w(\theta) = \theta$ for all θ; i.e. it solves the firm's problem:*

$$\begin{aligned} \max_{y, l_f(\theta)} \quad & y - \int w(\theta) l_f(\theta) dG(\theta) \\ s.t. \quad & y \leq \int \theta l_f(\theta) dG(\theta), \end{aligned}$$

3. *and, markets clear:*

   (a) $\int c(\theta) dG(\theta) = y$,

   (b) *and $l_f(\theta) = l(\theta)$, for all θ.*

1.1.1 Characterizing Equilibrium: General Case

Notice that utility is increasing in θ. That is, if we let $(c(\theta), l(\theta))$ denote the solution to the consumer's problem, and let $W(\theta) \equiv U(c(\theta), l(\theta))$, then, $W(\theta)$ is increasing in θ.

1.1.2 Adding Taxes and Transfers in the General Case

It is easy enough to add taxes and transfers to this model of the labor market. Given a labor tax rate, τ and a transfer, T, the household solves:

$$\begin{aligned} \max_{c,l} \quad & U(c,l) \\ s.t. \quad & c \leq (1-\tau) w(\theta) l + T. \end{aligned}$$

Have the Firm Problem be the same so that $w(\theta) = \theta$ for all θ and then add Government Budget Balance:

$\int \tau w(\theta)l(\theta)dG(\theta) = T = \int TdG(\theta).$

It follows immediately that it's still true that for any fixed τ, T, we have $W(\theta;\tau,T)$ is increasing in θ – Anything a low θ can do can also be chosen by a high θ but with extra consumption or leisure.

1.1.3 Special Case I: Log Utility

Suppose that $U(c,l) = \alpha \log c + (1-\alpha)\log(1-l)$. Then, consumer's problem is:

$$\begin{aligned} \max_{c,l} \quad & \alpha \log c + (1-\alpha)\log(1-l) \\ s.t. \quad & c \leq \theta l, \end{aligned}$$

which can be written as:

$$\max_{l} \ \alpha \log(\theta l) + (1-\alpha)\log(1-l).$$

First order condition is:

$$\begin{aligned} \frac{\alpha\theta}{\theta l} &= \frac{1-\alpha}{(1-l)}, \\ \therefore \frac{\alpha}{1-\alpha} &= \frac{l}{(1-l)}. \end{aligned}$$

Thus, $l(\theta)$ doesn't depend on θ. Indeed, $l = \alpha$ for all θ's. It follows that

$c(\theta) = \alpha\theta$

which is increasing in θ. Basically, the wealth of a θ-agent is θ and he splits this between c and leisure according to weights in utility as is standard with log utility.

Also, $W(\theta) = \alpha \log \theta + (1-\alpha)\log(1-\alpha)$, which is increasing in θ.

Proposition 2 *Those individuals who are* born *with low θ's would have been happier if they had been born with higher θ's.*

Proposition 3 *The Income Distribution, $F(y)$, is the distribution of $\alpha\theta$.*

$F(y) = \Pr(Y \leq y) = \Pr(\alpha\theta \leq y) = \Pr(\theta \leq y/\alpha) = G(y/\alpha).$

Adding Linear Income Taxes with Log Utility The new competitive equilibrium is the solution to:

$$\begin{aligned} \max_{c,l} \quad & \alpha \log c + (1-\alpha)\log(1-l) \\ s.t. \quad & c \leq (1-\tau)\theta l + T. \end{aligned}$$

Remark 4 *We need to add extra equilibrium condition here that the governments budget balances:*

$$\int \tau\theta l(\theta)\, dG(\theta) = \int T(\theta)\, dG(\theta) = T.$$

Consumer's problem can be written as:

$$\max_{l} \ \alpha \log((1-\tau)\theta l + T) + (1-\alpha)\log(1-l).$$

First order condition is:

$$\begin{aligned} \frac{\alpha(1-\tau)\theta}{[(1-\tau)\theta l + T]} &= \frac{1-\alpha}{(1-l)}, \\ \therefore \alpha(1-\tau)\theta(1-l) &= (1-\alpha)[(1-\tau)\theta l + T], \\ \therefore \alpha(1-\tau)(1-l) &= (1-\alpha)\left[(1-\tau)l + \frac{T}{\theta}\right], \qquad (1) \\ \therefore LHS(l;\theta) &\equiv \alpha(1-\tau)(1-l) \\ &= (1-\alpha)\left[(1-\tau)l + \frac{T}{\theta}\right] \\ &\equiv RHS(l;\theta). \end{aligned}$$

$LHS(l;\theta)$ is decreasing in l, as long as $0 < \tau < 1$. $RHS(l;\theta)$ is increasing in l as long as $0 < \tau < 1$. So there is a unique solution (if there is one). $LHS(l;\theta)$ doesn't change across θ's and $RHS(l;\theta)$ falls as θ goes up. Thus, $l(\theta)$ is increasing in θ if $T > 0$.

If $\tau > 0$ and $T = 0$, $l(\theta)$ is unchanged for all θ. Adding in $T > 0$ shifts $RHS(l;\theta)$ up for all θ's causing *all* $l(\theta)$'s to fall. How much depends on how big θ is. For example, if $\theta = \infty$, $l(\theta;\tau > 0, T = 0) = l(\theta;\tau > 0, T > 0)$.

Who is better off and who is worse off? For sure the net transfer satisfies:

$$T - \theta_{\max}\tau l(\theta_{\max}) \leq 0 \leq T - \theta_{\min}\tau l(\theta_{\min}),$$

since $l(\theta)$ is increasing in θ (see above). If everyone is the same (i.e., $G = \delta_{\{\theta\}}$ for some θ), then everyone is worse off (First Welfare Theorem) if $\tau > 0$.

If $\theta = 0$ is possible, they are are obviously helped. If θ is low enough, $l(\theta;\tau,T) = 0$. This will hold for all θ's such that the marginal value of leisure at $\ell = 1$ (so that $l = 0$)

exceeds the marginal utility of an additional unit of consumption, per unit of work.[2] The marginal value of leisure is $(1-\alpha)/(1-l)$, which at $l=0$ is $1-\alpha$, while the marginal value of extra consumption from working just a bit is

$$\frac{\alpha(1-\tau)\theta}{[(1-\tau)\theta l+T]},$$

which, when $l=0$, is $\alpha(1-\tau)\theta/T$; i.e. $l=0$ for all θ such that

$$\begin{aligned} 1-\alpha \geq & \frac{\alpha(1-\tau)\theta}{T}, \\ \text{or } \theta \leq & \frac{(1-\alpha)T}{\alpha(1-\tau)}. \end{aligned}$$

This is not as simple as it sounds, since T and τ are linked through the government's budget constraint.

Another way to see this is to rearrange Equation (1) to get: $l=\alpha-(1-\alpha)\frac{T}{\theta(1-\tau)}$.

Thus, in case of an interior solution, the labor supply would be an increasing function of θ, unlike the non-distorted case. However, even for the highest types, it would always be lower than the optimal non-distorted level α. Moreover, this is only true as long as $\theta \geq \frac{(1-\alpha)}{\alpha}\frac{T}{(1-\tau)}$.

Otherwise, it is optimal not to supply labor at all, $l(\theta)=0$. In this sense, this tax system makes very low types better off, by letting them exit the labor force, and still letting them eat T. This holds for all θ such that $T \geq \alpha\theta$. However, we still have to solve the model in more detail to see the overall effect.

Exercise 5 *Show how the equilibrium depends on the distribution of θ's. E.g. suppose* $\Pr(\theta=0.5)=\varepsilon$, *and* $\Pr(\theta=1.5)=1-\varepsilon$. *How does the equilibrium depend on ε? In particular,*

1. *When is it true that $l(0.5)=0$? Who is helped and hurt by a given (τ, T) scheme?*
2. *What does the Income Distribution look like in this case ($y(\theta)=\theta l(\theta)$ and $l(\theta)$ as functions of τ, T, and F)?*
3. *What about non-linear tax systems?*
4. *What if it is a disutility of work shock?*
5. *Can we show that the highest type is always made worse off?*

What if $\theta \sim U[0,2]$? What is the equilibrium? Who is made better off and who is made worse off?

[2] We use ℓ to denote leisure, while keep using l for the labor supply, henceforth.

1.1.4 Special Case II: Constant Elasticity of Substitution Utility

Consider the problem:

$$\begin{aligned} \max_{c,\ell} \quad & U(c,\ell) = \frac{a_1}{(1-\sigma)} c^{1-\sigma} + \frac{a_2}{(1-\sigma)} \ell^{1-\sigma} \\ s.t. \quad & c + \theta\ell \leq \theta, \end{aligned}$$

which can be written as:

$$\max_{\ell} \frac{a_1}{(1-\sigma)} \left[\theta(1-\ell)\right]^{1-\sigma} + \frac{a_2}{(1-\sigma)} \ell^{1-\sigma}.$$

First order condition implies:

$$\begin{aligned} a_1\theta^{1-\sigma}(1-\ell)^{-\sigma} =& a_2\ell^{-\sigma}, \\ \therefore a_1\theta^{1-\sigma}\left(\frac{\ell}{1-\ell}\right)^{\sigma} =& a_2, \\ \therefore \left(\frac{\ell}{1-\ell}\right)^{\sigma} =& \frac{a_2}{a_1\theta^{1-\sigma}}, \\ \therefore \frac{\ell}{1-\ell} =& \left(\frac{a_2}{a_1}\right)^{1/\sigma} \theta^{(\sigma-1)/\sigma}. \end{aligned}$$

Thus, $\frac{\ell}{1-\ell}$ and hence ℓ is decreasing in θ if and only if $\sigma < 1$; i.e. $l = 1-\ell$ is increasing in θ if and only if $\sigma < 1$. This is easy to see by looking at the two extremes, $\sigma = 0$ (perfect substitutes between c and ℓ and $\sigma = \infty$, perfect complements between c and ℓ).

1.2 What If There Were Insurance Against θ?

The simplest *Mirlees* example is when the government can see everyone's type, θ. There are two different, equivalent formulations:

1. government observes each person's θ, and tells him what to do, $(c(\theta), l(\theta))$; or,

2. government announces a tax schedule as a function of income and type, $T(y,\theta)$, and individual agents solve:

$$\begin{aligned} \max_{c,l} \quad & U(c,l) \\ s.t. \quad & c \leq \theta l - T(\theta l, \theta). \end{aligned}$$

Notice that:

1. any allocation of the form in Formulation 1, $(c(\theta), l(\theta))$, that satisfies

$$U(c(\theta), l(\theta)) \geq U(0, 1)$$

can be realized as an equilibrium with a tax schedule; just have the tax schedule give a θ-type 0 consumption and leisure if they don't work the right amount (i.e. to implement $(c^*(\theta), l^*(\theta))$, let $T(\theta l^*(\theta), \theta) = \theta l^*(\theta) - c^*(\theta)$ if $l = l^*$, and $T(\theta l, \theta) = \theta l(\theta)$ if $l \neq l^*$); and,

2. any equilibrium of the form in Formulation 2 can be realized as an order by the government; if the equilibrium is $(c^*(\theta), l^*(\theta))$, let this be what is dictated.

1.2.1 Benevolent Government with Full Information

Consider a *benevolent government* (or a *social planner*) that observes everyone's type, θ, and seeks to maximize an equally weighted utility of the agents by directly choosing the allocations (this pertains to Formulation 1 above). The problem of the government is:

$$\begin{aligned} \max_{\{c(\theta), l(\theta)\}} \quad & \int U(c(\theta), l(\theta))\, dG(\theta) \\ s.t. \quad & \int c(\theta)\, dG(\theta) \leq \int \theta l(\theta)\, dG(\theta). \end{aligned}$$

Consider the special case of separable utility function of the form $U(c, l) = u(c) - v(l)$, where, for usual preference properties, we need $u(\cdot)$ to be increasing and concave, and $v(\cdot)$ to be increasing and convex. Then, government's problem can be written as:

$$\begin{aligned} \max_{\{c(\theta), l(\theta)\}} \quad & \int [u(c(\theta)) - v(l(\theta))]\, dG(\theta) \\ s.t. \quad & \int c(\theta)\, dG(\theta) \leq \int \theta l(\theta)\, dG(\theta). \end{aligned}$$

Assume that there is a density for θ, $g(\theta) = dG(\theta)/d\theta$. Then, this becomes:

$$\begin{aligned} \max_{\{c(\theta), l(\theta)\}} \quad & \int [u(c(\theta)) - v(l(\theta))]\, g(\theta)\, d\theta \\ s.t. \quad & \int c(\theta)\, g(\theta)\, d\theta \leq \int \theta l(\theta)\, g(\theta)\, d\theta. \end{aligned}$$

First order conditions for this problem with respect to

1. $c(\theta)$ is $u'(c(\theta))\, g(\theta) = \lambda g(\theta)$, and
2. $l(\theta)$ is $v'(l(\theta))\, g(\theta) = \lambda \theta g(\theta)$.

Therefore, $u'(c(\theta)) = \lambda$ for all θ, and, as a result, $c(\theta) = \bar{c}$ for all θ, where $u'(\bar{c}) = \lambda$. Moreover, $v'(l(\theta)) = \lambda\theta$. Hence,

$$\begin{aligned} v'(l(\theta)) =& u'(c(\theta))\,\theta \\ =& u'(\bar{c})\,\theta. \end{aligned} \tag{2}$$

This means, $c(\theta)$ doesn't depend on θ, while $l(\theta)$ is increasing in θ. Thus, $U(c(\theta), 1 - l(\theta))$ is decreasing in θ.

As a result, the solution to the benevolent government's problem is not *incentive compatible,* if θ is only privately observed; i.e. if θ was not observable for the government, and it wanted to implement the allocation derived above, individuals would not have incentives to reveal their types truthfully, if asked to do so.

Implementing the Optimum with Taxes What if we wanted to implement the solution to the government's problem in the previous section, through a tax system? (This pertains to Formulation 2 above.) The consumer's problem, in this case, is:

$$\begin{aligned} \max_{c,l} \quad & U(c, l) \\ s.t. \quad & c \leq \theta l - T(\theta l, \theta), \end{aligned}$$

which can be written as:

$$\max_{l} \; U(\theta l - T(\theta l, \theta), l).$$

The first order condition is:

$$\theta U_1(\theta l - T(\theta l, \theta), l)\,[1 - T_1(\theta l, \theta)] = -U_2(\theta l - T(\theta l, \theta), l).$$

For our (separable) utility function, this becomes:

$$\theta u'(c(\theta))\,[1 - T_1(\theta l(\theta), \theta)] = v'(l(\theta)), \quad \forall \theta.$$

where $(c(\theta), l(\theta))$ is the solution to an individual of type-θ.

From the first order conditions of the planner's problem above (Equation (2)), $v'(l(\theta)) = u'(c(\theta))\,\theta$ for all θ. Thus, to implement this with a (smooth) tax system, it must be true that

$$1 - T_1(\theta l(\theta), \theta) = 1, \quad \forall \theta;$$

i.e. $T_1(\theta l(\theta), \theta) = 0$ for all θ. This means, all taxes are lump sum. However, these lump sum amounts are type specific, $T(y, \theta) = T(\theta)$.

Since everyone has the same consumption, $c(\theta) = \bar{c}$, and higher types have higher labor supply, $l(\theta)$, it follows that $T(\theta)$ is decreasing in θ.

Exercise 6 *What is income distribution? [With log utility, high types have higher pre-tax income (they work more), all types have the same after tax income (since it is given by c).] What if $U(\cdot)$ is not separable? Who is happy, who is not?*

Insurance Why is this called insurance? Suppose that we gave each person, ex ante, a choice between the government's *contract* and the ex post competitive equilibrium allocation. Ex ante, every agent would choose to go with the government's contract; it insures agents against the future realization of their type. Indeed, you could rephrase the government's contracting problem as one of a profit maximizing insurance company. You'd get exactly the same first order conditions if you also imposed a minimum expected utility on the contracts that they could offer.

1.2.2 Private Information Version

Assume that only the agent can see θ and hours, $l(\theta)$, but that the government can see the output, $\theta l(\theta) = y(\theta)$, and consumption, $c(\theta)$.

Again, there are two formulations:

1. the government announces a contract $(c(\theta), l(\theta))$, but is restricted in that it can only choose contracts where it is in the agents' own self-interest (given that θ is private info) to deliver on those contracts; or,

2. the government announces a tax function, $T(y)$, and then agents solve

$$\begin{aligned} \max_{c,l} \quad & U(c, 1-l) \\ s.t. \quad & c \leq \theta l - T(\theta l), \end{aligned}$$

   to get $(c(\theta), l(\theta))$, given their type, θ; i.e. T can no longer depend on θ.

Contract Formulation The problem of a benevolent government (or, a social planner) that seeks to maximize an equally weighted average of the utilities, in the case of separable utility form, is now given by:

$$\begin{aligned} \max_{\{c(\theta),l(\theta)\}} \quad & \int [u(c(\theta)) - v(l(\theta))]\, g(\theta)\, d\theta \\ s.t. \quad & \int c(\theta)\, g(\theta)\, d\theta \leq \int \theta l(\theta)\, g(\theta)\, d\theta \\ & u(c(\theta)) - v(l(\theta)) \geq u\left(c\left(\hat{\theta}\right)\right) - v\left(\frac{\hat{\theta}}{\theta} l\left(\hat{\theta}\right)\right), \quad \forall \theta, \hat{\theta}. \end{aligned}$$

The first constraint is Feasibility (FEAS) and the second is Incentive Compatability (IC).

This means output, $\theta l(\theta)$, is directly observable for each person, but neither θ nor $l(\theta)$ is. Thus, a person of productivity θ can *pretend* to be a person of type $\hat{\theta}$ by producing the same output as is required of a person of type $\hat{\theta}$. This requires $\frac{\hat{\theta}}{\theta} l\left(\hat{\theta}\right)$ from a person of type θ.

Note: This is not the formulation that Mirrlees originally used in his paper on the topic. Rather, he formulated the problem as choosing an income tax function, $T(y)$, and an allocation, $(c(\theta), l(\theta))$ subject to the constraint that:

$$(c(\theta), l(\theta)) \text{ solves } \max_{(c,l)} \quad u(c) - v(l) \text{ subject to } c \leq \theta l - T(\theta l).$$

This last statement is the equivalent of the 'Implementability Constraint' from the Ramsey Problems we have studied in the past.

These two versions of the problem – the 'Contracting' version and the TDCE one are equivalent. That is, Incentive Compatability in the problem as stated above is the 'Implementability Constraint' for the Mirrlees problem.

Problem: Show that these two versions of the problem are equivalent.

As a side note that the ex post Competitive Equilibrium allocation satisfies (IC). This is also true of all of the TDCE allocations discussed above for any τ and T, or even general $T(y)$ income tax functions, but, the ex ante full information allocation does not.

1.2.3 Simple Version #1

Assume that there are exactly two types, θ_H and θ_L, with $\theta_H > \theta_L$, $\Pr(\theta = \theta_H) = \pi_H$, and $\Pr(\theta = \theta_L) = \pi_L$. Then, the planning problem becomes:

$$\max_{c_L, l_L, c_H, l_H} \quad \pi_H [u(c_H) - v(l_H)] + \pi_L [u(c_L) - v(l_L)] \tag{SP1}$$

$$s.t. \quad \pi_H c_H + \pi_L c_L \leq \pi_H \theta_H l_H + \pi_L \theta_L l_L \tag{FEAS}$$

$$u(c_H) - v(l_H) \geq u(c_L) - v\left(\frac{\theta_L}{\theta_H} l_L\right) \tag{IC1}$$

$$u(c_L) - v(l_L) \geq u(c_H) - v\left(\frac{\theta_H}{\theta_L} l_H\right), \tag{IC2}$$

where Constraint (FEAS) is the feasibility condition, Constraint (IC1) ensures that it is in the θ_H-type agents' (high-type agents) own interest to deliver on their contract (i.e. to reveal themselves truthfully), and Constraint (IC2) ensures this is the case for the θ_L-type agents. Constraints (IC1) and Constraint (IC2) are referred to as *incentive compatibility* constraints for the high and low-types, respectively.

Exercise 7 *Show that (FEAS) holds with equality.*

To characterize the solution to Problem (SP1), we proceed in the following steps.

Step 1: Rewrite the Problem Rewrite the problem in terms of output requirements by types. Let y_H be the output required of a high type and y_L that required of a low type. Then, Problem (SP1) becomes:

$$\max_{c_L, y_L, c_H, y_H} \quad \pi_H \left[u(c_H) - v\left(\frac{y_H}{\theta_H}\right) \right] + \pi_L \left[u(c_L) - v\left(\frac{y_L}{\theta_L}\right) \right] \tag{SP2}$$

$$s.t. \quad \pi_H c_H + \pi_L c_L \leq \pi_H y_H + \pi_L y_L \tag{FEAS}$$

$$u(c_H) - v\left(\frac{y_H}{\theta_H}\right) \geq u(c_L) - v\left(\frac{y_L}{\theta_H}\right) \tag{IC1}$$

$$u(c_L) - v\left(\frac{y_L}{\theta_L}\right) \geq u(c_H) - v\left(\frac{y_H}{\theta_L}\right). \tag{IC2}$$

Note: This problem is typically not concave/convex – concave objective function and a convex constraint set. This is because the $u(c)'s$ (and $v's$) appear on both the left and right hand side of the (IC) constraints.

Step 2: Note Some Properties of Feasible Contracts Here we note some simple properties of contracts, i.e. combinations of (c_L, y_L) and (c_H, y_H) that must be true if (FEAS), (IC1), and (IC2) are all satisfied.

1. Suppose $c_H > c_L$ but $y_H \leq y_L$. If this were true, then,

$$u(c_H) - v\left(\frac{y_H}{\theta_L}\right) > u(c_L) - v\left(\frac{y_H}{\theta_L}\right),$$

   since $c_H > c_L$ and $u(\cdot)$ is monotone. Moreover,

$$u(c_L) - v\left(\frac{y_H}{\theta_L}\right) \geq u(c_L) - v\left(\frac{y_L}{\theta_L}\right),$$

   since $y_L \geq y_H$ and $v(\cdot)$ is monotone. Thus,

$$u(c_H) - v\left(\frac{y_H}{\theta_L}\right) > u(c_L) - v\left(\frac{y_L}{\theta_L}\right).$$

   But this violates (IC2) and hence, these types of allocations are not feasible.

2. A similar argument shows that combinations with $c_H \geq c_L$ and $y_H < y_L$ also are not feasible.

3. Suppose $c_H < c_L$ but $y_L \leq y_H$. If this were true, as above, we would have

$$u(c_L) - v\left(\frac{y_L}{\theta_H}\right) > u(c_H) - v\left(\frac{y_H}{\theta_H}\right);$$

   i.e. (IC1) would be violated.

4. A similar argument holds if $c_H \leq c_L$ but $y_L < y_H$. So, this cannot be the case in the solution.

We can summarize these in the following lemma.

Lemma 8 *If the contract* (c_L, y_L) *and* (c_H, y_H) *satisfies* (FEAS), (IC1), *and* (IC2) *in Problem* (SP2), *then, one of the following three configurations must hold (Note - FEAS isn't really needed for these):*

1. $c_H > c_L$ *and* $y_H > y_L$; *OR,*
2. $c_L > c_H$ *and* $y_L > y_H$; *OR,*
3. $c_L = c_H$ *and* $y_L = y_H$.

Step 3: Showing that Configuration 2 is Not Feasible (Note - FEAS is not used here either.) First some simple intuition: note that the indifference curves of the high type are, at every point (in $c-y$ space) steeper than those of the low type; fix a pair, (c^*, y^*). At this point, the slope of an indifference curve of a type θ agent is given by (note that indifference curves are upward sloping in this space, as in Figure 1):

$$\begin{aligned}
\left.\frac{dy}{dc}\right|_{U(c^*,y^*;\theta)} &= -\left.\frac{\partial U\left(c,y;\theta\right)/\partial c}{\partial U\left(c,y;\theta\right)/\partial y}\right|_{(c,y)=(c^*,y^*)} \\
&= -\left.\frac{u'\left(c\right)}{-\frac{1}{\theta}v'\left(\frac{y}{\theta}\right)}\right|_{(c,y)=(c^*,y^*)} \\
&= \theta\frac{u'\left(c^*\right)}{v'\left(\frac{y^*}{\theta}\right)}.
\end{aligned}$$

Since $v\left(\cdot\right)$ is convex and y/θ is decreasing in θ, $v'\left(y/\theta\right)$ is decreasing in θ holding y fixed (at y^*). Thus, increasing θ increases $\frac{1}{v'(y/\theta)}$ and increases θ. As a result, increasing θ increases $\theta\frac{1}{v'(y/\theta)}u'(c) = \frac{dy}{dc}$ at any point (c, y). In summary, higher θ types have steeper indifference curves than lower θ types through any (c, y) pair.

The reason that this matters is as follows: suppose that it is true that $(c_L, y_L) > (c_H, y_H)$. Consider the indifference curve of the low-type that passes through the point (c_H, y_H). By (IC2), this corresponds to a (weakly) lower level of utility to the low-type. But, since the indifference curves of the high-type are everywhere steeper than those of the low-type, it follows that the indifference curve of the high-type through the point (c_H, y_H) passes *above* (c_L, y_L). This implies that the high type prefers the choice (c_L, y_L); (IC1) is violated. This is shown graphically in Figure 2. Thus, it must be true that either $(c_L, y_L) < (c_H, y_H)$ or $(c_L, y_L) = (c_H, y_H)$. In other words, the optimal contract is monotone increasing in types.

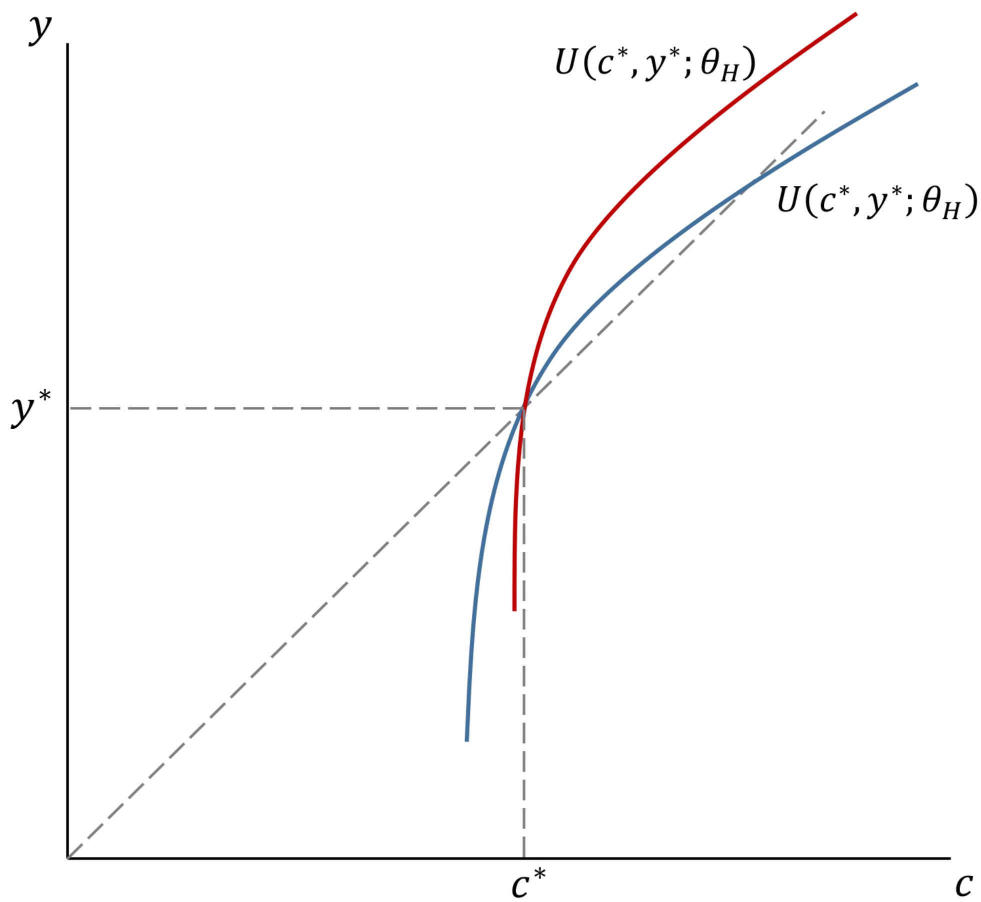

Figure 1: Indifference Curves of the Two Types

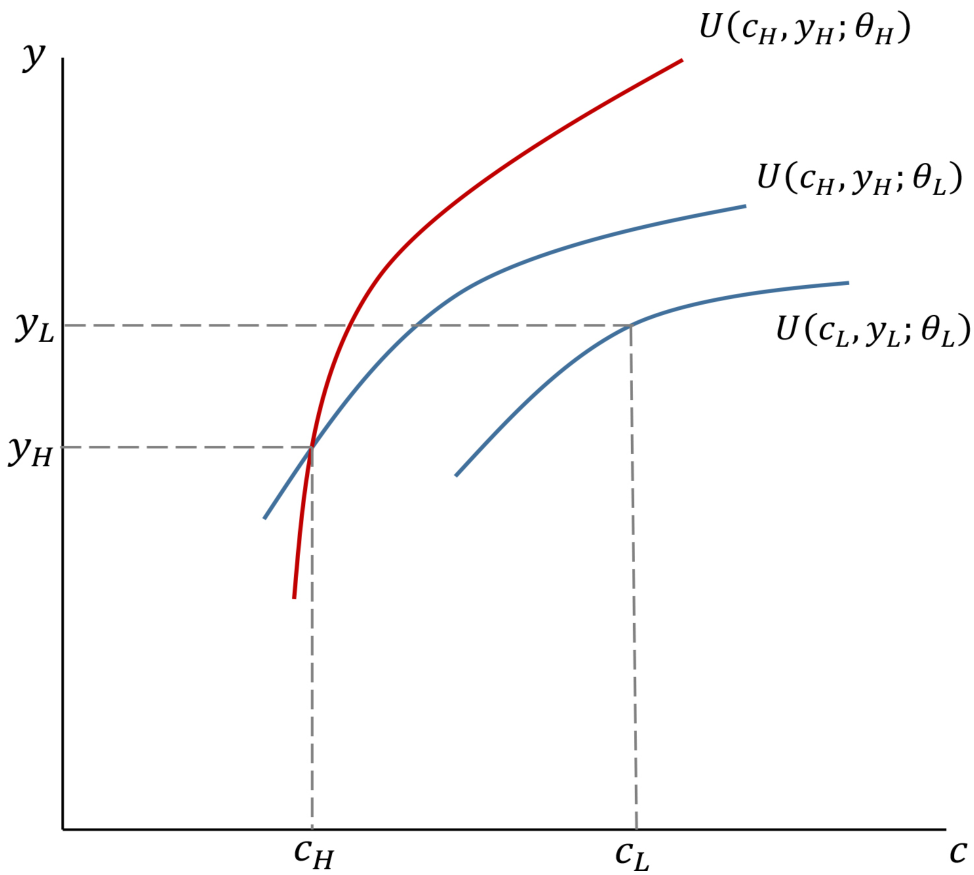

Figure 2: Monotonicity of Optimal Contracts in Types

To show this formally, assume $(c_L, y_L) > (c_H, y_H)$. By (IC2),

$$U(c_L, y_L; \theta_L) - U(c_H, y_H; \theta_L) = u(c_L) - u(c_H) - \left[v\left(\frac{y_L}{\theta_L}\right) - v\left(\frac{y_H}{\theta_L}\right)\right] \geq 0,$$

or,

$$\begin{aligned} u(c_L) - u(c_H) - \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy &\geq 0, \\ \therefore u(c_L) - u(c_H) &\geq \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy. \end{aligned}$$

Notice that the first term, $u(c_L) - u(c_H)$ is positive, since c_L is assumed to be larger than c_H. Also, since $v'(\cdot) > 0$ and $y_L > y_H$, it follows that the second term is also positive. But since $\theta_H > \theta_L$ and $v(\cdot)$ is convex, it follows that $v'(y/\theta_H) < v'(y/\theta_L)$ for all y, and, as a result,

$$\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy < \int_{y_H}^{y_L} v\left(\frac{y}{\theta_L}\right) dy.$$

Since $\theta_H > \theta_L$, we also have:

$$\frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy < \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy.$$

Thus,

$$\begin{aligned} u(c_L) - u(c_H) \geq & \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy \\ > & \frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy, \end{aligned}$$

implying that,

$$\begin{aligned} u(c_L) - u(c_H) - \frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy > & 0 \\ \therefore U(c_L, y_L; \theta_H) - U(c_H, y_H; \theta_H) = & u(c_L) - u(c_H) - \left[v\left(\frac{y_L}{\theta_H}\right) - v\left(\frac{y_H}{\theta_H}\right)\right] > 0. \end{aligned}$$

That is high-type agents prefer (c_L, y_L) over (c_H, y_H), violating (IC1).

From this, we conclude that either $c_L = c_H$ and $y_L = y_H$, or $c_H > c_L$ and $y_H > y_L$.

(This is the analog of the single crossing property.)

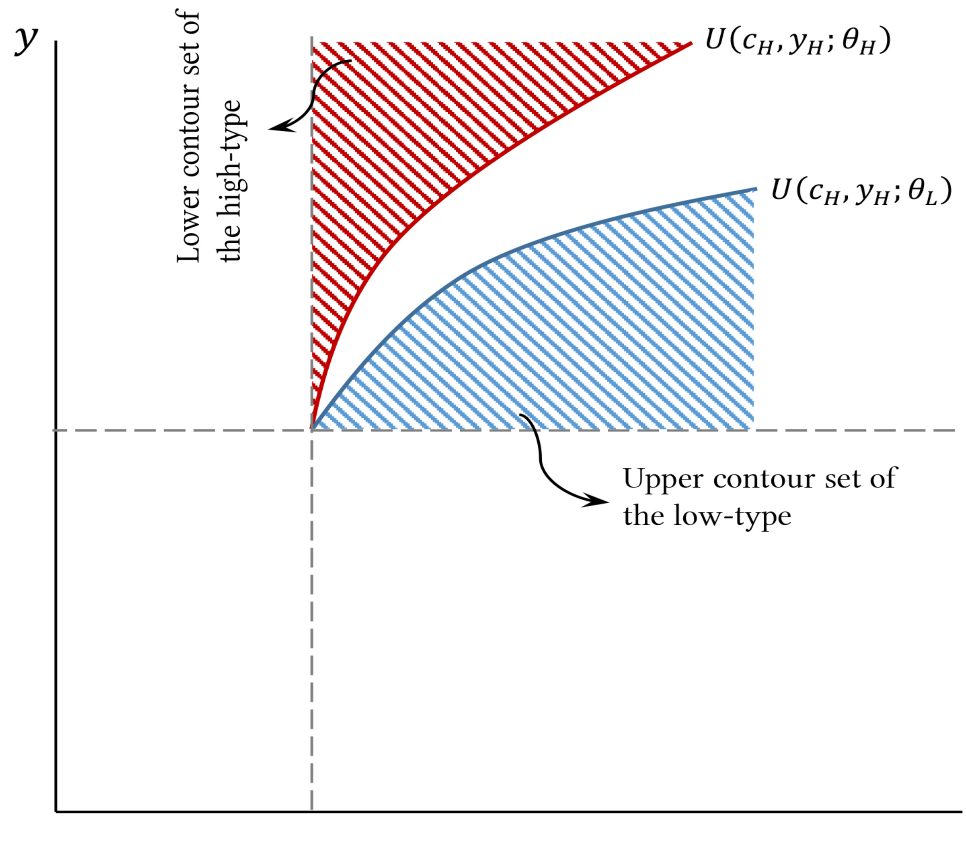

Figure 3: High-Type Contract as the Intersection of Upper Contour Sets

Another way to think of Configuration 2 above is to consider the *contour sets* of the preferences; consider the bundle (c_H, y_H) and the upper contour set of the low-types, and the lower contour set of the high-types passing through this point. The intersections of these sets with the set

$$\{(c, y) \mid c \geq c_H,\ y \geq y_H\}$$

do not intersect with each other, other than (c_H, y_H), as shown in Figure 3.

Step 4: Showing that Configuration 3 is Not Optimal Next, we want to claim that $c_H = c_L$ and $y_H = y_L$ cannot be a solution to the planning problem in (SP2).

To provide some intuition why this is the case, first of all, consider an autarkic allocation in which each agent is left alone to consumer whatever he produces; that is consumption-production choice of an agent of type θ is the solution to:

$$\max_{c,y} \quad u(c) - v\left(\frac{y}{\theta}\right)$$
$$s.t. \quad c \le y,$$

which can be written as:

$$\max_{y} \ u(y) - v\left(\frac{y}{\theta}\right).$$

The first order condition for this problem implies that:

$$u'(c(\theta)) = \frac{1}{\theta} v'\left(\frac{y(\theta)}{\theta}\right),$$

for each type θ. As θ increases, $c(\theta)$ must increase for this equation to hold. Otherwise, the left hand side increases, while all the terms on the right hand side decrease. As a result, higher types would be consuming and producing more in the autarkic allocations. This result holds strictly, if $u(\cdot)$ and $v(\cdot)$ satisfy the usual assumptions. Notice also that this autarkic allocation is feasible, in the sense that it satisfies Conditions (FEAS), (IC1), and (IC2) in the planning problem in (SP2).

On the other hand, when $c_H = c_L$ and $y_H = y_L$, different types are again, in effect, in autarky; they consume whatever they produce. The question is, when the planner distorts the decisions by forcing different types by producing differently than they would have done if left alone, would this result in an increase in the average utility? The intuitive answer is no; without providing any insurance or redistribution (note that, agents consume what they produce), this only has distortionary effects.

To show this formally, suppose (c, y) denotes the common consumption/production pair. We consider the following cases:

1. If
$$u'(c) < \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right),$$
since $\theta_H > \theta_L$, we have:
$$u'(c) < \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right).$$
Therefore, decreasing c and y at the same time by a small amount will keep the (IC)s holding and will be strictly better for both types.

2. If
$$u'(c) = \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right),$$
again, we have:
$$u'(c) < \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right).$$

Hence, decreasing consumption and production of the low types would make them better off, while high types have no incentives to deviate to the new allocation.

3. If
$$u'(c) > \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right) \text{ and } u'(c) > \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right),$$
by the same argument as in the first case, it is optimal to increase y and c at the same time. The case of
$$u'(c) > \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right) \text{ and } u'(c) = \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right),$$
is not optimal either, by the same logic as in the second case.

4. If
$$u'(c) > \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right) \text{ and } u'(c) < \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right),$$
an increase in the consumption of high types followed by an increase in their production, and a decrease in the consumption of low types followed by a decrease in their production would leave both types better off.

Lemma 9 *Monotonicity. At the optimal allocation, $c_H > c_L$ and $y_H > y_L$.*

Step 5: Showing that We Can Drop (IC2) The planner's problem can now be written as:

$$\max_{c_L, y_L, c_H, y_H} \quad \pi_H \left[u(c_H) - v\left(\frac{y_H}{\theta_H}\right)\right] + \pi_L \left[u(c_L) - v\left(\frac{y_L}{\theta_L}\right)\right] \tag{SP3}$$
$$s.t. \quad \pi_H c_H + \pi_L c_L \leq \pi_H y_H + \pi_L y_L \tag{FEAS}$$
$$u(c_H) - v\left(\frac{y_H}{\theta_H}\right) \geq u(c_L) - v\left(\frac{y_L}{\theta_H}\right) \tag{IC1}$$
$$u(c_L) - v\left(\frac{y_L}{\theta_L}\right) \geq u(c_H) - v\left(\frac{y_H}{\theta_L}\right) \tag{IC2}$$
$$c_L < c_H \text{ and } y_L < y_H. \tag{MONOT}$$

Note: MONOT is a redundant constraint given the last Lemma.

Consider the following relaxed version of this maximization problem:

$$\max_{c_L, y_L, c_H, y_H} \quad \pi_H \left[u(c_H) - v\left(\frac{y_H}{\theta_H}\right)\right] + \pi_L \left[u(c_L) - v\left(\frac{y_L}{\theta_L}\right)\right] \tag{RP1}$$

$$s.t. \quad \pi_H c_H + \pi_L c_L \le \pi_H y_H + \pi_L y_L \tag{FEAS}$$

$$u(c_H) - v\left(\frac{y_H}{\theta_H}\right) \ge u(c_L) - v\left(\frac{y_L}{\theta_H}\right) \tag{IC1}$$

$$c_L < c_H \text{ and } y_L < y_H. \tag{MONOT}$$

Since there are strictly less constraints for this problem, it follows that if a solution exists, and if that solution satisfies (IC2), it is a solution for Problem (SP3), too. We will show that this is the case.

Suppose that the solution to (RP1) is (c_L, y_L) and (c_H, y_H). We'll show that (IC1) is satisfied at equality. We will do this by supposing it is false and constructing a better contract. The dominating contract that we will construct will have better insurance over c without disrupting (IC1).

Suppose that (IC1) is *not* satisfied at equality;

$$u(c_H) - v\left(\frac{y_H}{\theta_H}\right) > u(c_L) - v\left(\frac{y_L}{\theta_H}\right).$$

Notice that, if this holds, by continuity, it will still hold if we add a bit to c_L and subtract a bit from c_H;

$$u(c_H - \varepsilon) - v\left(\frac{y_H}{\theta_H}\right) > u(c_L + \delta) - v\left(\frac{y_L}{\theta_H}\right),$$

as long as ε and δ are small enough.

Consider the alternative contract given by $(c_H - \varepsilon, y_H)$ and $(c_L + \delta, y_L)$. Choose $\delta = \pi_H \varepsilon / \pi_L$. Then, if ε is small enough, (IC1) will still hold, and (FEAS) becomes:

$$\begin{aligned} \pi_H (c_H - \varepsilon) + \pi_L (c_L + \delta) =& \pi_H c_H + \pi_L c_L + \pi_H \varepsilon - \pi_L \frac{\pi_H}{\pi_L} \varepsilon \\ =& \pi_H c_H + \pi_L c_L. \end{aligned}$$

Thus, (FEAS) will hold because we didn't change y_L or y_H, and because of the way we constructed δ. So, we only need to show that welfare goes up from this change, even when ε is small but positive. To see this note that the change in welfare is given by:

$$\Delta W = \pi_H [u(c_H - \varepsilon) - u(c_H)] + \pi_L \left[u\left(c_L + \frac{\pi_H}{\pi_L}\varepsilon\right) - u(c_L)\right].$$

The terms involving the y's do not appear in this, since they are unchanged. If we take the derivative with respect to ε, at $\varepsilon = 0$, we have:

$$\begin{aligned}\left.\frac{d\Delta W}{d\varepsilon}\right|_{\varepsilon=0} &= -\pi_H u'(c_H) + \pi_L \frac{\pi_H}{\pi_L} u'(c_L) \\ &= -\pi_H u'(c_H) + \pi_H u'(c_L) \\ &= \pi_H \left[u'(c_L) - u'(c_H)\right] \\ &> 0,\end{aligned}$$

since $c_L < c_H$, and $u(\cdot)$ is assumed to be strictly concave.

Thus, (RP1) is equivalent to:

$$\begin{aligned} \max_{c_L, y_L, c_H, y_H} \quad & \pi_H \left[u(c_H) - v\left(\frac{y_H}{\theta_H}\right)\right] + \pi_L \left[u(c_L) - v\left(\frac{y_L}{\theta_L}\right)\right] && \text{(RP2)} \\ s.t. \quad & \pi_H c_H + \pi_L c_L \leq \pi_H y_H + \pi_L y_L && \text{(FEAS)} \\ & u(c_H) - v\left(\frac{y_H}{\theta_H}\right) = u(c_L) - v\left(\frac{y_L}{\theta_H}\right) && \text{(IC1)} \\ & c_L < c_H \text{ and } y_L < y_H. && \text{(MONOT)} \end{aligned}$$

Lastly, to show that (IC2) is redundant in Problem (SP3), it suffices to show that at the solution to (RP1), (IC2) is satisfied; i.e. we want to show that, if

$$u(c_H) - v\left(\frac{y_H}{\theta_H}\right) = u(c_L) - v\left(\frac{y_L}{\theta_H}\right),$$

and

$$c_L < c_H \text{ and } y_H < y_L,$$

then

$$u(c_L) - v\left(\frac{y_L}{\theta_L}\right) > u(c_H) - v\left(\frac{y_H}{\theta_L}\right).$$

This follows from the fact that the indifference curves of the low type are flatter than those of the high type; Figure 4 illustrates this point graphically; given (c_H, y_H), by the above argument (that (IC1) holds with equality at the solution), (c_L, y_L) must lie on the same indifference curve of the high-types passing through (c_H, y_H), $U(c_H, y_H; \theta_H)$. Let $U(c_H, y_H; \theta_L)$ be the indifference curve of the low-types passing through this point. Any allocation lying above $U(c_H, y_H; \theta_L)$ would mean that low-types would prefer (c_H, y_H). Anything bellow $U(c_H, y_H; \theta_L)$ imply that the low types weakly prefer the new allocation to (c_H, y_H). However, when $(c_H, y_H) \neq (c_L, y_L)$, the new allocation is strictly preferred to (c_H, y_H).

We can summarize all of this in the following proposition.

Proposition 10 *The solution to* (RP2) *is the same as the solution to* (SP1)*. The optimal allocation is monotone in type, (IC1) holds with equality and (IC2) is slack.*

See the Appendix to see the generalization of this result to the multiple, finite, type case.

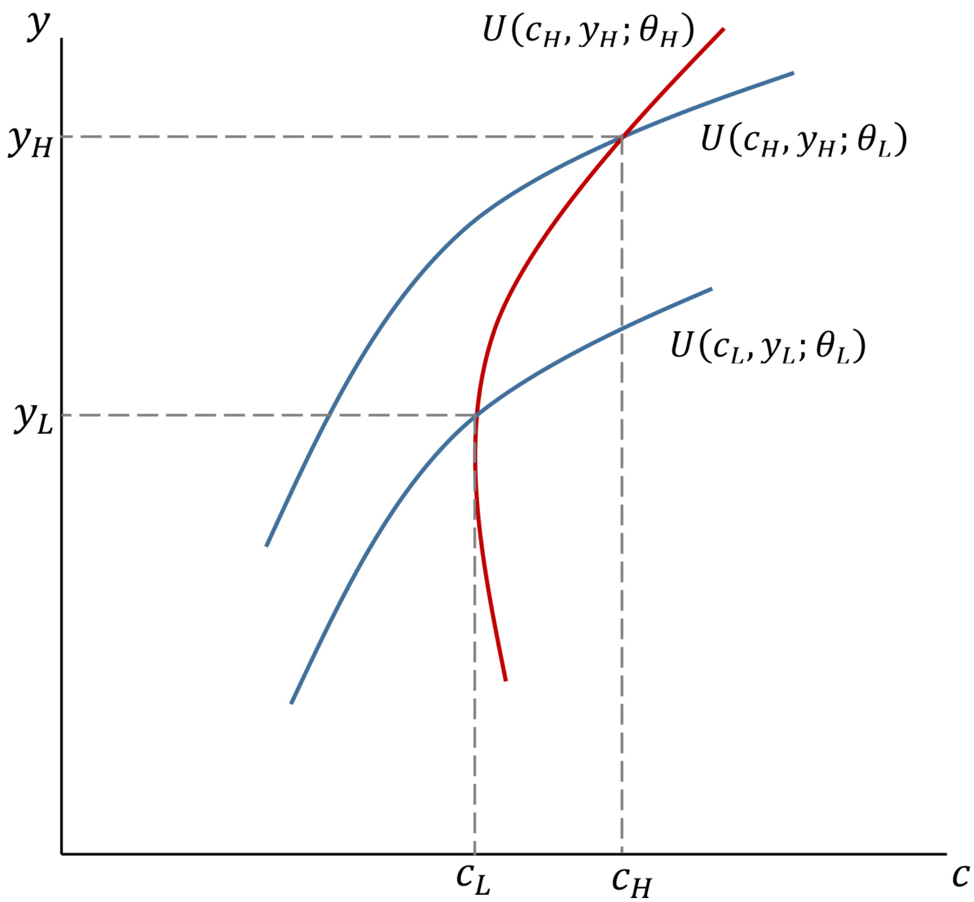

Figure 4: Indifference of the High-Types between Two Contracts

1.3 Characterizing the Solution

1.3.1 Average Tax Rates

As a first step in characterizing the solution let's start with something simple. Who subsidizes who? I.e., we know that the high type produces and consumes more than the low type, but, does he produce more than he consumes? Or less? To answer this question, we'll first need a result concerning marginal tax rates.

1.3.2 Marginal Tax Rates

By Proposition 10, the planning problem can be written as:

$$\max_{c_L,l_L,c_H,l_H} \quad \pi_H\left[u\left(c_H\right)-v\left(l_H\right)\right]+\pi_L\left[u\left(c_L\right)-v\left(l_L\right)\right] \tag{SP4}$$

$$s.t. \quad \pi_H c_H+\pi_L c_L \leq \pi_H\theta_H l_H+\pi_L\theta_L l_L \tag{FEAS}$$

$$u\left(c_H\right)-v\left(l_H\right)=u\left(c_L\right)-v\left(\frac{\theta_L}{\theta_H}l_L\right) \tag{IC1}$$

If we let λ and μ be the Lagrange multipliers on (FEAS) and (IC1), respectively, the Lagrangian can be written as:

$$\begin{aligned}\mathcal{L}=\pi_H\left[u\left(c_H\right)-v\left(l_H\right)\right]+\pi_L\left[u\left(c_L\right)-v\left(l_L\right)\right]&\\ +\lambda\left(\pi_H\theta_H l_H+\pi_L\theta_L l_L-\pi_H c_H-\pi_L c_L\right)&\\ +\mu\left[u\left(c_H\right)-v\left(l_H\right)-u\left(c_L\right)+v\left(\frac{\theta_L}{\theta_H}l_L\right)\right]&.\end{aligned}$$

First order conditions are:

1. c_H: $\pi_H u'\left(c_H\right)+\mu u'\left(c_H\right)=\pi_H\lambda$;
2. l_H: $\pi_H v'\left(l_H\right)+\mu v'\left(l_H\right)=\pi_H\theta_H\lambda$;
3. c_L: $\pi_L u'\left(c_L\right)-\mu u'\left(c_L\right)=\pi_L\lambda$; and,
4. l_L: $\pi_L v'\left(l_L\right)-\mu\frac{\theta_L}{\theta_H}v'\left(\frac{\theta_L}{\theta_H}l_L\right)=\pi_L\theta_L\lambda$.

From Conditions 1 and 2, we get:

$$\frac{\left(\pi_H+\mu\right)v'\left(l_H\right)}{\left(\pi_H+\mu\right)u'\left(c_H\right)}=\frac{\pi_H\theta_H\lambda}{\pi_H\lambda},$$

or,

$$\begin{aligned}\frac{v'\left(l_H\right)}{u'\left(c_H\right)}&=\theta_H,\\ \therefore\frac{v'\left(l_H\right)}{\theta_H}&=u'\left(c_H\right).\end{aligned} \tag{3}$$

This is the same condition that we get in the *full information* case.

From Conditions 3 and 4, we get:

$$\frac{\pi_L v'(l_L) - \mu \frac{\theta_L}{\theta_H} v'\left(\frac{\theta_L}{\theta_H} l_L\right)}{(\pi_L - \mu) u'(c_L)} = \frac{\pi_L \theta_L \lambda}{\pi_L \lambda},$$

or,

$$\begin{aligned} \frac{\pi_L v'(l_L) - \mu \frac{\theta_L}{\theta_H} v'\left(\frac{\theta_L}{\theta_H} l_L\right)}{(\pi_L - \mu) u'(c_L)} &= \theta_L, \\ \therefore \frac{\left[\pi_L - \mu \frac{\theta_L}{\theta_H} v'\left(\frac{\theta_L}{\theta_H} l_L\right) / v'(l_L)\right]}{(\pi_L - \mu)} v'(l_L) &= \theta_L u'(c_L). \end{aligned} \tag{4}$$

Next, we show that $\mu < \pi_L$, and that $\frac{\theta_L}{\theta_H} v'\left(\frac{\theta_L}{\theta_H} l_L\right) / v'(l_L) < 1$. To see the first part, use Conditions 1 and 3 to write:

$$\begin{aligned} \frac{(\pi_H + \mu) u'(c_H)}{(\pi_L - \mu) u'(c_L)} &= \frac{\pi_H}{\pi_L}, \\ \therefore (\pi_H + \mu) u'(c_H) &= \frac{\pi_H}{\pi_L} (\pi_L - \mu) u'(c_L). \end{aligned}$$

Since $(\pi_H + \mu) > 0$, $u'(c_H) > 0$, $\frac{\pi_H}{\pi_L} > 0$, and $u'(c_L) > 0$, it follows that $\mu < \pi_L$.

Next, if $\theta_L < \theta_H$, since $v'(\cdot)$ is strictly increasing, we have

$$v'\left(\frac{\theta_L}{\theta_H} l_L\right) < v'(l_L),$$

which implies that $v'\left(\frac{\theta_L}{\theta_H} l_L\right) / v'(l_L) < 1$, and, hence,

$$\begin{aligned} \frac{\theta_L}{\theta_H} \frac{v'\left(\frac{\theta_L}{\theta_H} l_L\right)}{v'(l_L)} &< 1, \\ \therefore \mu \frac{\theta_L}{\theta_H} \frac{v'\left(\frac{\theta_L}{\theta_H} l_L\right)}{v'(l_L)} &< \mu. \end{aligned}$$

From this it follows that

$$\frac{\left[\pi_L - \mu \frac{\theta_L}{\theta_H} \frac{v'\left(\frac{\theta_L}{\theta_H} l_L\right)}{v'(l_L)}\right]}{(\pi_L - \mu)} > 1,$$

since $\mu < \pi_L$, as shown above.

Thus, from Equation (4), we have:

$$\begin{aligned} v'(l_L) =& \frac{(\pi_L - \mu)}{\left[\pi_L - \mu \frac{\theta_L}{\theta_H} \frac{v'\left(\frac{\theta_L}{\theta_H} l_L\right)}{v'(l_L)}\right]} \theta_L u'(c_L) \\ <& \theta_L u'(c_L), \end{aligned}$$

or,

$$\frac{v'(l_L)}{\theta_L} < u'(c_L). \tag{5}$$

This means the value of leisure to the low-type agents is less than it's productivity at the margin. If we define τ so that

$$\frac{v'(l_L)}{(1-\tau)\theta_L} = u'(c_L),$$

then we must have $0 < \tau < 1$.

From these results, Equality (3) and Inequality (5), comes the standard intuition about optimal contracts; the high-types are undistorted at the margin, but the low-types are.

Remark 11 *If $\frac{\theta_L}{\theta_H} = 1$, i.e. there is no private information, then,*

$$\begin{aligned} v'(l_L) =& \frac{(\pi_L - \mu)}{\left[\pi_L - \mu \frac{\theta_L}{\theta_H} \frac{v'\left(\frac{\theta_L}{\theta_H} l_L\right)}{v'(l_L)}\right]} \theta_L u'(c_L) \\ =& \frac{(\pi_L - \mu)}{(\pi_L - \mu)} \theta_L u'(c_L) \\ =& \theta_L u'(c_L), \end{aligned}$$

and, as a result, the low-type is also undistorted.

1.3.3 Average Tax Rates Again

We'll show that average taxes are negative for the low type and positive for the high type:

Lemma 12 *In the optimal allocation:*

$$y_H - c_H > 0 > y_L - c_L.$$

Proof. Recall that (IC1) holds with equality, so we can rewrite it as:

$$u(c_H) - u(c_L) = v(\frac{y_H}{\theta_H}) - v(\frac{y_L}{\theta_H}).$$

Or,

$$\int_{c_L}^{c_H} u'(c)dc = \int_{y_L}^{y_H} \frac{1}{\theta_H} v'(\frac{y}{\theta_H})dy.$$

Now, because u is concave, we know that

$$u'(c_H)(c_H - c_L) < \int_{c_L}^{c_H} u'(c)dc < u'(c_L)(c_H - c_L).$$

By the Intermediate Value Theorem then, we have:

$$\int_{c_L}^{c_H} u'(c)dc = u'(c^*)(c_H - c_L) \quad for\ some\ c^* \in (c_L, c_H).$$

Similarly,

$$\int_{y_L}^{y_H} \frac{1}{\theta_H} v'(\frac{y}{\theta_H})\ dy = (y_H - y_L)\frac{1}{\theta_H} v'(\frac{y^*}{\theta_H})\ for\ some\ y^* \in (y_L, y_H).$$

Hence,

$$(c_H - c_L)u'(c^*) = (y_H - y_L)\frac{1}{\theta_H} v'\left(\frac{y^*}{\theta_H}\right).$$

Now, since u is concave and v is convex, we have:

$$u'(c^*) > u'(c_H)\ and\ \frac{1}{\theta_H} v'(\frac{y^*}{\theta_H}) < \frac{1}{\theta_H} v'(\frac{y_H}{\theta_H}).$$

So,

$$u'(c^*) > u'(c_H) = \frac{1}{\theta_H} v'\left(\frac{y_H}{\theta_H}\right) > \frac{1}{\theta_H} v'(\frac{y^*}{\theta_H}).$$

Where we have used the result from above that

$$\frac{1}{\theta_H} v'\left(\frac{y_H}{\theta_H}\right) = u'\left(c_H\right).$$

Thus,

$$(c_H - c_L) < (y_H - y_L).$$

This can be rewritten as:

$$(y_H - c_H) > (y_L - c_L).$$

Finally, from (FEAS) it can't true

$$y_H > c_H \; and. y_L > c_L$$

OR

$$y_H < c_H \; and \; y_L < c_L.$$

So, we must have that

$$y_H - c_H > 0 > y_L - c_L \; i.e., \; y_H > c_H \; and \; y_L < c_L.$$

■

1.3.4 Implementing the Optimal Contract with Taxes

Next, we discuss how to implement the contractual outcome, characterized above, through decentralized decisions by workers subject to income taxes. That is, we want to find an income tax schedule, $T(y)$, such that for each θ_i, the contractual allocation, (c_i, y_i), is the solution to:

$$\begin{aligned} \max_{c,y} \quad & u(c) - v\left(\frac{y}{\theta_i}\right) \\ s.t. \quad & c \leq y - T(y). \end{aligned} \tag{6}$$

In general there is no unique way of doing this; there are many different functions $T(y)$ such that will implement the optimal contractual outcome. For example, suppose $T(y) = y$ for all y other than y_L and y_H, and that $T(y_L) = y_L - c_L$, $T(y_H) = y_H - c_H$. If this is $T(y)$, it follows immediately that no one would ever choose any level of output other than y_L or y_H, since they would get $c = 0$. Thus, the only relevant options given this tax scheme are (c_L, y_L) and (c_H, y_H). Then, from (FEAS) and (IC1), it follows that the low-type chooses (c_L, y_L), and the high-type chooses (c_H, y_H).

A second alternative is to choose $T(y)$ so that the mapping $y \mapsto y - T(y)$ follows the indifference curve of the low-type for y's below y_L, and follows that of the high-type for y's above y_L; i.e.

1. for $y \leq y_L$, $u(y - T(y)) - v\left(\frac{y}{\theta_L}\right) = u(y_L - T(y_L)) - v\left(\frac{y_L}{\theta_L}\right)$; and,
2. for $y \geq y_H$, $u(y - T(y)) - v\left(\frac{y}{\theta_H}\right) = u(y_H - Ty_H) - v\left(\frac{y_H}{\theta_H}\right)$.

This makes the low-type indifferent between picking any $y \leq y_L$, and the high type indifferent between picking any $y \geq y_L$. Moreover, given the characterization of the contract above, it follows that the low-type is strictly worse off by picking any $y > y_L$, and the high-type is strictly worse off by picking any $y < y_L$.

Let us assume that the effective tax schedule for type i takes the form $T_i(y_i) = a_i + \tau_i y_i$, at least near the optimal contractual solution, y_i. We want to examine the properties of this tax schedule.

To this end, consider the first order conditions of the household problem, (6), given λ the Lagrange multiplier on the budget constraint:

1. c: $u'(c) = \lambda$; and,
2. y: $\frac{1}{\theta}v'\left(\frac{y}{\theta}\right) = \lambda\left[1 - T'(y)\right]$.

Therefore, for the intra-temporal condition on output and consumption, we have:

$$\frac{1}{\theta}v'\left(\frac{y}{\theta}\right) = u'(c)\left[1 - T'(y)\right]. \tag{7}$$

If we compare Condition (7) with (3), we observe that, to implement the optimum, planner must set

$$T'_H\left(y_{\theta_H}\right) = 0,$$

implying that

$$\tau_H = 0.$$

In other words, the tax schedule for the high types must look like a lump-sum taxation/transfer at the optimal level of output; as mentioned before, no distortion in the margin at the top (of the type distribution).

Next, comparing (7) with (5) reveals that

$$0 < \tau_L < 1.$$

As a result, the margin must be distorted for the low types. Moreover, this distortion is towards less work, and higher leisure.

Before asking ourselves why would the planner want the low-types' labor-consumption choice to be distorted toward working less, let us characterize the tax schedule further. In particular, so far, we know that it should look like $T(y) = a_\theta + \tau_\theta y$, with $\tau_H = 0$ and $\tau_L \in (0,1)$. What can be said about a_i, for the two types?

To answer this question, , we'll go back to our average tax question first.

We know from above that $a_H > 0$, since the high type consumes less than his output and faces a zero marginal tax rate. As a result, we must have $a_H > 0$ near y_H. Then, for the resulting allocation to be optimal, the fact that feasibility must hold with equality at the optimum, implies that $a_L < 0$ near y_L.

In summary, the optimal tax schedule near (c_H, y_H) is exactly similar to a lump-sum taxation, whereas, near (c_L, y_L), tax schedule looks like a linear tax on output (income), and a lump-sum rebate. Hence, we can summarize this *special case* as a two-part tax schedule, where

1'. $T^*(y) = -T_L + T'(y_L)\, y_L$, for all $y \leq y_L$ (an affine tax function, with a non-zero intercept), and

2'. $T^*(y) = T_H$, for all $y \geq y_L$ (a lump sum tax function),

in which T_L, T_H, and $T'(y_L)$ are strictly positive. As shown in Figure 5, in this two-part income tax scheme, each party would pick its component of the optimal contract. It has the advantage that each component is linear, and they differ in their lump sum transfers.

The expression for $\tau_L = T'(y_L)$ was given above.

To find T_L and T_H, we can use the budget constraints:

$$T_L = c_L - (1-\tau_L)y_L > 0 \; from \; c_L > y_L > (1-\tau_L)y_L$$

and

$$T_H = y_H - c_H > 0.$$

Now, we are ready to answer the question that why the planner wants to tax output for the low-types, creating a distortion toward lower output; subsidizing the low-types creates an incentive for the high-types to pretend they are of low-type. To prevent this from happening, planner must keep the low-types output-consumption schedule on the indifference curve of the high-types (as illustrated by Step 5). Planner does this by putting a linear tax on income near y_L. In fact, this linear tax is more meant to deter high-types from lying.

Labor Supply Implications We know that $y_H > y_L$ in the optimum. But, can we tell who works more? In other words, is it also true that $l_H > l_L$?

This is not an easy problem, and not much is known about it. In what follows, we will use what we did above concerning the marginal tax rates for the two types along with the conclusion that there is a transfer from the high type to the low type.

Let us assume that the tax scheme looks like the one described in the previous section; a two-part tax schedule that is linear near the optimal contractual optima. Define the labor supply function, $l\left((1-\tau)\, w, T\right)$, as the solution to

$$\begin{aligned} \max_{c,l} \quad & u(c) - v(l) \\ s.t. \quad & c \leq (1-\tau)\, wl + T. \end{aligned}$$

Note that leisure is a normal good in this problem, and, hence, $l\left((1-\tau)\, w, T\right)$ is strictly decreasing in T, holding w and τ constant. As a result,

$$l_H = y_H/\theta_H = l\left((1-\tau)\, w = \theta_H, T_H < 0\right) > l\left(\theta_H, 0\right);$$

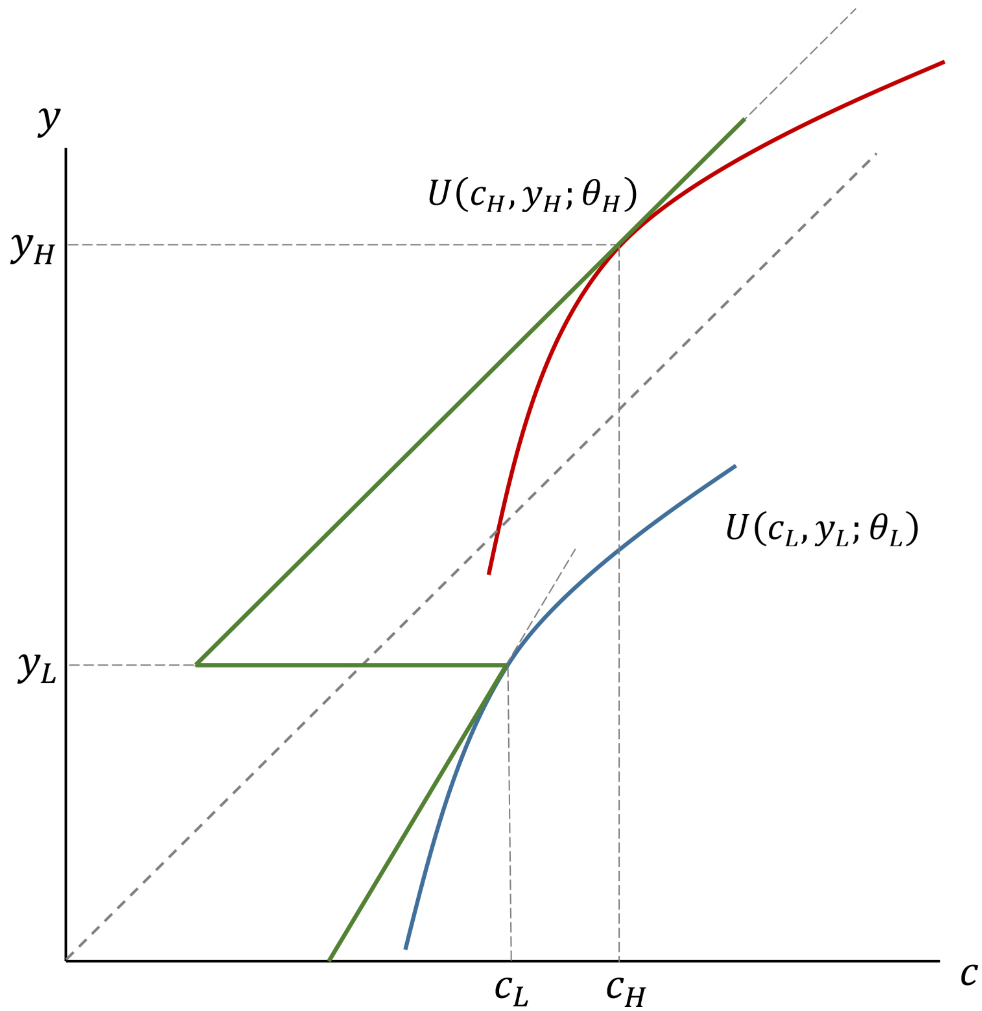

Figure 5: Two-Part Optimal Tax Scheme

i.e., since the high type has a zero marginal tax rate and faces a negative transfer, he works more than he would without taxes (in autarky).

Let's make the Assumption that the labor supply function is upward sloping (weakly) in net wages; i.e., $l((1-\tau)w, T)$ is increasing in $(1-\tau)w$, holding T fixed.

Under this assumption, it follows that

$$l_L = l\left(\left[1 - T'\left(y_L\right)\right]\theta_L, T_L\right) < l\left(\theta_L, T_L\right) < l\left(\theta_H, T_L\right) < l\left(\theta_H, 0\right),$$

where the last step comes from the fact that $T_L > 0$, given that leisure is a normal good. As a result, under the Assumption above, it follows that

$$l_L < l\left(\theta_H, 0\right) < l\left(\theta_H, T_H\right) = l_H$$

since leisure is a normal good and $T_H < 0$.

How do l_H and l_L compare to what they were in the ex ante competitive equilibrium? Here, I will use the superscript ce to denote the allocation in the ex post competitive equilibrium allocation discussed at the beginning of these notes.

First,.leisure is a normal good in this model due to the separability of preferences. Thus, for the high type, since the marginal tax rate is zero and the transfer is negative, it follows that $l_H > l_H^{ce}$. For the low type, the transfer is positive, lowering labor supply. In addition to that, the after tax wage is also lower which also decreases labor supply. It follows that $l_L < l_L^{ce}$.

Problem: Can you show how c_L and c_H compare to c_L^{ce} and c_H^{ce}? I didn't try to do this...

Log Utility Log-utility is an example in which Assumption **??** is satisfied, and, hence, labor supplied by the low-types is lower than that of the high-types. Intuitively, *in the ex post competitive equilibrium*, under log-utility, income and substitution effects cancel out, such that $l_H = l_L$ in autarky, while $c_H = \theta_H l_H > \theta_L l_L$. Planner desires to redistribute from high-type agents to the low-types. As noted above, he does this by levying a positive lump-sum tax on the high-type. This shifts the equilibrium budget constraint in the consumption-leisure space, (c, ℓ), inward in a parallel way. Thus, the high-types consume less leisure and less consumption good than without the redistribution, resulting to an increase in their labor supply. This revenue is rebated to the low-type agents and a labor income tax is also levied against them. In the end, this twists the budget constraint of the low-type to make it steeper (since the labor income tax is positive for this type, as shown above), but also shifts it out. These two act in the direction of increasing the consumption of leisure in the low-types; a decrease in their labor supply, as in Figure 7.

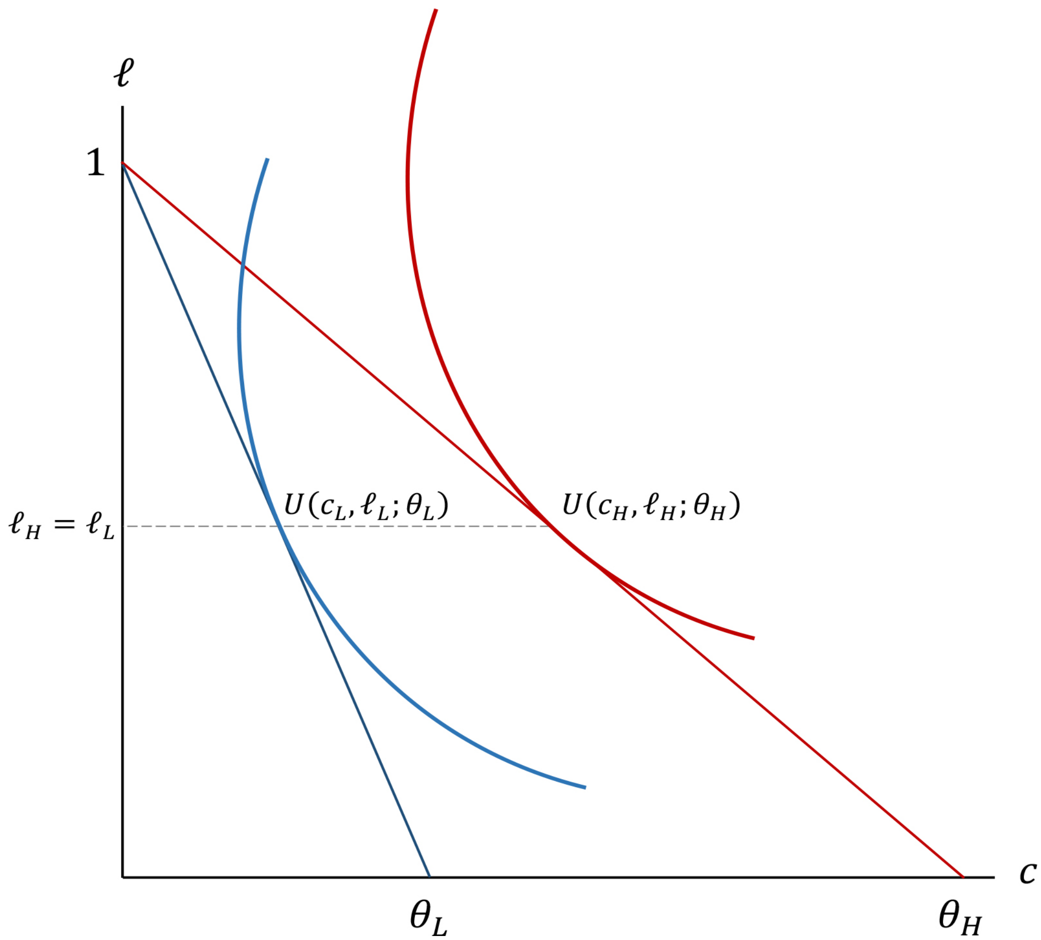

Figure 6: Optimal Leisure-Consumption Choice in Autarky

Thus, the high-type consumes less leisure than the competitive equilibrium, while the low-type consumes more. Since they consume the same amount at the competitive equilibrium, it follows that at the Mirrleesian allocation, $l_H > l_L$.

In what follows we give a formal proof of this result by using the implementation of the Mirrlees allocation using taxes, using the fact that $T_H < 0$ and $T_L > 0$, i.e. the planner moves resources from the high-type to the low-type.

Assume that $U(c, l) = \alpha \log(c) + (1 - \alpha) \log(1 - l)$. In the Competitive Equilibrium allocation without insurance we have the following allocation (we have normalized the price of the consumption good to $p_c = 1$):

$$\begin{aligned} c_H^{ce} &= \alpha\theta_H, \quad \ell_H^{ce} = 1 - \alpha, \quad l_H^{ce} = \alpha, \\ c_L^{ce} &= \alpha\theta_L, \quad \ell_L^{ce} = 1 - \alpha, \text{ and } l_L^{ce} = \alpha. \end{aligned}$$

The important thing to note is that $l_L^{ce} = l_H^{ce} = \alpha$, and $\ell_H^{ce} = \ell_L^{ce} = 1 - \alpha$.

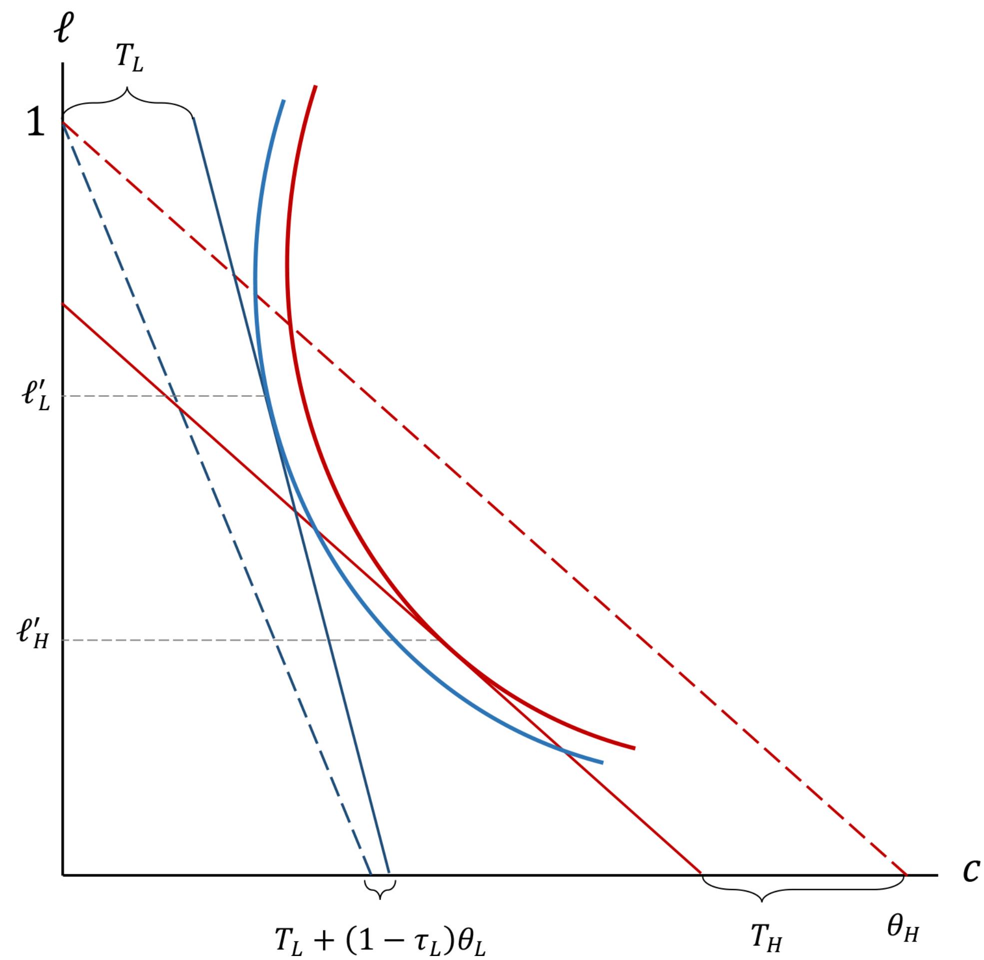

Figure 7: Optimal Leisure-Consumption Choice with Taxes

Next, consider the optimal decision of an agent of type s, when faced with an income tax rate of τ_s and a lump sum transfer of T_s:

$$\begin{aligned} \max_{c,l} \quad & \alpha \log (c) + (1-\alpha) \log (1-l) \\ s.t. \quad & c \leq (1-\tau_s)\, \theta_s l + T_s, \end{aligned}$$

which can be written as

$$\max_{c,l} \quad \alpha \log ((1-\tau_s)\, \theta_s l + T_s) + (1-\alpha) \log (1-l)\,.$$

Necessary conditions for the solution are

$$\begin{aligned} \frac{\alpha (1-\tau_s)\, \theta_s}{c_s} &= \frac{1-\alpha}{1-l_s}, \\ \text{and } c_s &= (1-\tau_s)\, \theta_s l_s + T_s. \end{aligned}$$

We can rearrange the first equality as

$$c_s = \frac{\alpha}{1-\alpha} (1-\tau_s)\, \theta_s (1-l_s)\,.$$

And, therefore, in the optimum, we must have:

$$(1-\tau_s)\, \theta_s l_s + T_s = \frac{\alpha}{1-\alpha} (1-\tau_s)\, \theta_s (1-l_s)\,.$$

Dividing by $(1-\tau_s)\, \theta_s$, we can write this equality as

$$l_s = \frac{\alpha}{1-\alpha} (1-l_s) - \frac{T_s}{(1-\tau_s)\, \theta_s},$$

which can be rearranged as

$$l_s \left(\frac{1-\alpha+\alpha}{1-\alpha} \right) = \frac{\alpha}{1-\alpha} - \frac{T_s}{(1-\tau_s)\, \theta_s},$$

implying that

$$l_s = \alpha - \frac{(1-\alpha)\, T_s}{(1-\tau_s)\, \theta_s},$$

in the optimum.

Note that, if $\tau_s = 0$ and $T_s < 0$, it follows that $l_s > \alpha$, while, if $T_s > 0$, then $l_s < \alpha$, and it is decreasing in τ_s, as long as $\tau_s < 1$. Thus,

$$l_H^{mir} > \alpha > l_L^{mir},$$

as desired.

Another Special Case This section looks to be a dead end and nothing was done with it. It is just here for now for possible future use!

Suppose $v(l) = a_2 l^{\alpha}$, with $\alpha > 1$. Then, $v'(l) = \alpha a_2 l^{\alpha-1} \equiv x$, and

$$l = (v')^{-1}(x) = \left(\frac{x}{\alpha a_2}\right)^{\frac{1}{(\alpha-1)}}.$$

In this case, the first order condition for the low-types implies

$$\begin{aligned} l_L &= \left[\left(\frac{\theta_H}{\theta_L}\right)^{\alpha-1}\right]^{\frac{1}{(\alpha-1)}} \left[\frac{1}{\alpha a_2}\frac{(\pi_L-\mu)}{\left(\pi_L-\mu\frac{\theta_L}{\theta_H}\right)}\theta_L u'(c_L)\right]^{\frac{1}{(\alpha-1)}} \\ &= \left[\frac{1}{\alpha a_2}\left(\frac{\theta_H}{\theta_L}\right)^{\alpha-1}\frac{(\pi_L-\mu)}{\left(\pi_L-\mu\frac{\theta_L}{\theta_H}\right)}\theta_L u'(c_L)\right]^{\frac{1}{(\alpha-1)}}. \end{aligned}$$

$$\begin{aligned} u'(c_H) &> \frac{\theta_L}{\theta_H}\left(\frac{\theta_H}{\theta_L}\right)^{\alpha-1}\frac{(\pi_L-\mu)}{\left(\pi_L-\mu\frac{\theta_L}{\theta_H}\right)}u'(c_L) \\ \therefore u'(c_H) &> \left(\frac{\theta_L}{\theta_H}\right)^{1+(1-\alpha)}\frac{(\pi_L-\mu)}{\left(\pi_L-\mu\frac{\theta_L}{\theta_H}\right)}u'(c_L). \end{aligned}$$

First order condition for the high-types implies

$$\begin{aligned} l_H &= (v')^{-1}[\theta_H u'(c_H)] \\ &= \left[\frac{1}{\alpha a_2}\theta_H u'(c_H)\right]^{\frac{1}{(\alpha-1)}}. \end{aligned}$$

Thus, $l_H > l_L$ if, and only if,

$$\left[\frac{1}{\alpha a_2}\theta_H u'(c_H)\right]^{\frac{1}{(\alpha-1)}} > \left[\frac{1}{\alpha a_2}\left(\frac{\theta_H}{\theta_L}\right)^{\alpha-1}\frac{(\pi_L-\mu)}{\left(\pi_L-\mu\frac{\theta_L}{\theta_H}\right)}\theta_L u'(c_L)\right]^{\frac{1}{(\alpha-1)}}.$$

It is not clear if it is worth it to do this case any further!?

1.3.5 Generalizations

Most of these arguments seem pretty general; the solution is monotone increasing in type, and no one ever pretends to be of higher type. However, we still need to show that only locally downward incentive constraints bind (i.e. θ_i does not pretend to be

θ_{i-2}, if $\theta_j \geq \theta_{j-1}$ for all $j \geq 2$). This is shown in Appendix 3 (that has been added by Anderson).

The exception to this is the details of showing that $l_H > l_L$. The current argument for this depends heavily on the assumption of log-utility. We have already shown that much of this argument could be extended, at least in the two type case, to situations in which the labor supply curve is upward sloping (Assumption **??** is satisfied);

1. $l_H^{mirr} < l_H^{ce}$, since leisure is a normal good, and the high-type becomes poorer when $T_H > 0$;

2. since leisure is normal, $T_L > 0$, the labor supply curve slopes up, and $(1 - \tau_L)\,\theta_L < \theta_L$, we have that $l_L^{mirr} < l_L^{ce}$;

3. since the labor supply curve is upward sloping, and

$$(1 - \tau_L)\,\theta_L < \theta_H = (1 - \tau_H)\,\theta_H,$$

   we have that $l_L^{ce} < l_H^{ce}$;

4. and, finally, $l_L^{mirr} < l_L^{ce} < l_H^{ce} < l_H^{mirr}$, as desired.

In terms of what kinds of utility functions satisfy Assumption **??**, suppose

$$u\,(c, \ell) = \frac{a_1}{(1 - \sigma)} c^{1-\sigma} + \frac{a_2}{(1 - \sigma)} \ell^{1-\sigma} \text{ with } \sigma \geq 0.$$

Then the labor supply curve slopes up if, and only if, $0 \leq \sigma < 1$. Thus, this corresponds to the low-curvature case; not the most natural assumption! See Appendix 3 for a discussion of this.

1.4 Implementation in Contracts

Could either of these allocations (full Info or private information) be implemented by private firms offering wage contracts (with incentive compatibility constraints, where relevant)? For the answer to be positive, agents require the ability to sign perfectly enforceable contracts *before* the realization of the shock. Is this possible in the real world?

2 Mirrlees with a Continuum of Types

(These notes come from Ali Shourideh.)

Suppose there is a continuum of types, distributed according to the distribution F, and the planner wants to maximized the average of utilities weighted by H. So, the planning problem can be written as:

$$\begin{aligned}\max_{c(\theta),y(\theta)} \quad & \int \left[u\left(c\left(\theta\right)\right) - v\left(\frac{y\left(\theta\right)}{\theta}\right)\right] dH\left(\theta\right) \\ s.t. \quad & \int \left[c\left(\theta\right) - y\left(\theta\right)\right] dF\left(\theta\right) \leq 0 \\ & \theta \in \arg\max_{\hat{\theta}} \left\{ u\left(c\left(\hat{\theta}\right)\right) - v\left(\frac{y\left(\hat{\theta}\right)}{\theta}\right)\right\}.\end{aligned}$$

In what follows, we are mostly interested in utilitarian planner, where $H = F$.

The major difficulty in dealing with this problem, in the vastness and complexity of the incentive constraint; it represents a continuum of constraints for each type θ. If we could replace this by a single stationarity constraint for each type, it would made out lives way easier.

One way to do so is to proceed as we did in the previous section; show that, in the optimum, only the locally downward incentive constraints bind. Then, we need only to care about local deviations of types, and a constraint that ensures the local stationarity of the solution would do the job; a constraint like the following:

$$u'\left(c\left(\theta\right)\right) c'\left(\theta\right) - v'\left(\frac{y\left(\theta\right)}{\theta}\right)\frac{y'\left(\theta\right)}{\theta} = 0. \tag{8}$$

Recall our steps in the previous section; after showing that consumption and output are both increasing in types in the optimum, we proved, in Lemmas 8 and 12 that only downward incentive constraint binds in the optimum. This enabled us to simplify the problem greatly, and make some inference regarding the solution. Why can't we follow a same procedure here?

Assume we could somehow prove that the output is increasing in types, $y'\left(\theta\right) > 0$. Clearly, consumption would be increasing in types as well, $c'\left(\theta\right) > 0$. By the exact same logic as in Lemmas 8 and 12, only locally downward incentive constraints bind in the optimum. Therefore, replacing the incentive constraints by the stationarity condition of (8) would not affect the results. This is widely known as the first order approach in the literature.

Let us assume for the moment that $y'\left(\theta\right) > 0$, and write the relaxed planning problem as:

$$\begin{aligned}\max_{c(\theta),y(\theta)} \quad & \int \left[u\left(c\left(\theta\right)\right) - v\left(\frac{y\left(\theta\right)}{\theta}\right)\right] dH\left(\theta\right) \\ s.t. \quad & \int \left[c\left(\theta\right) - y\left(\theta\right)\right] dF\left(\theta\right) \leq 0 \\ & u'\left(c\left(\theta\right)\right) c'\left(\theta\right) - v'\left(\frac{y\left(\theta\right)}{\theta}\right)\frac{y'\left(\theta\right)}{\theta} = 0, \quad \forall\theta.\end{aligned}$$

Our approach is to solve this relaxed version of the model, and observe its implications regarding the optimal output. If $y'(\theta) > 0$ in the solution, then the first order approach is indeed valid.

To solve the problem, define

$$U(\theta) \equiv u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right).$$

Then, by the stationarity condition, we must have:

$$U(\theta) = \max_{\hat{\theta}} \; u\left(c\left(\hat{\theta}\right)\right) - v\left(\frac{y\left(\hat{\theta}\right)}{\theta}\right).$$

Therefore:

$$U'(\theta) = \frac{y(\theta)}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right).$$

Thus, we may write the problem with $U(\theta)$ as a control as:

$$\begin{aligned}
\max_{U(\theta), c(\theta), y(\theta)} \quad & \int U(\theta)\, dH(\theta) \\
s.t. \quad & U(\theta) = u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right), \quad \forall \theta \\
& \int [c(\theta) - y(\theta)]\, dF(\theta) \leq 0 \\
& U'(\theta) = \frac{y(\theta)}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right), \quad \forall \theta.
\end{aligned}$$

The Lagrangian for this problem is:

$$\begin{aligned}
\mathcal{L} = & \int U(\theta)\, dH(\theta) + \int \gamma(\theta) \left[u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right) - U(\theta)\right] d\theta \\
& + \lambda \int [y(\theta) - c(\theta)]\, dF(\theta) + \int \mu(\theta) \left[U'(\theta) - \frac{y(\theta)}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right)\right] d\theta.
\end{aligned}$$

Using the integration by parts, we may write the last term as:

$$\int \mu(\theta) U'(\theta)\, d\theta = \mu\left(\bar{\theta}\right) U\left(\bar{\theta}\right) - \mu(\underline{\theta}) U(\underline{\theta}) - \int \mu'(\theta) U(\theta)\, d\theta.$$

Therefore, the Lagrangian becomes:

$$\begin{aligned}
\mathcal{L} = & \int U(\theta)\, dH(\theta) + \int \gamma(\theta) \left[u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right) - U(\theta)\right] d\theta \\
& + \lambda \int [y(\theta) - c(\theta)]\, dF(\theta) - \int \mu(\theta) \frac{y(\theta)}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right) d\theta \\
& + \mu\left(\bar{\theta}\right) U\left(\bar{\theta}\right) - \mu(\underline{\theta}) U(\underline{\theta}) - \int \mu'(\theta) U(\theta)\, d\theta.
\end{aligned}$$

Given the Lagrange multipliers at the solution, the objective is to solve the unconstrained problem of maximizing $\mathcal{L}$. If $\mu\left(\bar{\theta}\right) \neq 0$ at the solution, then the optimal thing to do is to set $c\left(\bar{\theta}\right)$ to infinity (by giving measure one of output to a measure zero of types) and $y\left(\bar{\theta}\right) = 0$, as to tend $U\left(\bar{\theta}\right)$ to infinity (or negative infinity if $\mu\left(\bar{\theta}\right) < 0$). However, this would certainly violate the incentive constraints of the other types. Similarly, $\mu\left(\underline{\theta}\right) \neq 0$ cannot be the case in the solution. Therefore, we can write this Lagrangian as

$$\begin{aligned}\mathcal{L} = & \int U(\theta)\, dH(\theta) + \int \gamma(\theta) \left[u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right) - U(\theta) \right] d\theta \\ & + \lambda \int \left[y(\theta) - c(\theta) \right] dF(\theta) - \int \mu(\theta) \frac{y(\theta)}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right) d\theta - \int \mu'(\theta)\, U(\theta)\, d\theta,\end{aligned}$$

with $\mu\left(\bar{\theta}\right) = \mu\left(\underline{\theta}\right) = 0$, as the boundary conditions.

Now that, since everything in the Lagrangian is in integral form, we may simply take the first order conditions, without caring about the integrals. These conditions are:

1. $h(\theta) - \gamma(\theta) - \mu'(\theta) = 0$;
2. $\gamma(\theta)\, u'(c(\theta)) - \lambda f(\theta) = 0$; and,
3. $-\gamma(\theta) \frac{1}{\theta} v'\left(\frac{y(\theta)}{\theta}\right) + \lambda f(\theta) - \mu(\theta) \left[\frac{y(\theta)}{\theta^3} v''\left(\frac{y(\theta)}{\theta}\right) + \frac{1}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right) \right] = 0.$

By the first condition, we have:

$$\begin{aligned}\mu\left(\bar{\theta}\right) - \mu(\theta) &= \int_{\theta}^{\bar{\theta}} \mu'\left(\hat{\theta}\right) d\hat{\theta} \\ &= \int_{\theta}^{\bar{\theta}} \left[h\left(\hat{\theta}\right) - \gamma\left(\hat{\theta}\right) \right] d\hat{\theta}.\end{aligned}$$

Therefore, noting that $\mu\left(\bar{\theta}\right) = \mu\left(\underline{\theta}\right) = 0$, we have:

$$\mu(\theta) = \int_{\theta}^{\bar{\theta}} \left[\gamma\left(\hat{\theta}\right) - h\left(\hat{\theta}\right) \right] d\hat{\theta}.$$

By the second condition, then:

$$\mu(\theta) = \int_{\theta}^{\bar{\theta}} \left[\frac{\lambda f\left(\hat{\theta}\right)}{u'\left(c\left(\hat{\theta}\right)\right)} - h\left(\hat{\theta}\right) \right] d\hat{\theta}.$$

By substituting this into Condition 3, and using Condition 2 again, we get:

$$\lambda f(\theta) - \frac{\lambda f(\theta)}{u'(c(\theta))}\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right)$$
$$= \left[\int_{\theta}^{\bar{\theta}} \left[\frac{\lambda f\left(\hat{\theta}\right)}{u'\left(c\left(\hat{\theta}\right)\right)} - h\left(\hat{\theta}\right)\right] d\hat{\theta}\right] \left[\frac{y(\theta)}{\theta^3}v''\left(\frac{y(\theta)}{\theta}\right) + \frac{1}{\theta^2}v'\left(\frac{y(\theta)}{\theta}\right)\right],$$

Therefore:

$$1 - \frac{1}{u'(c(\theta))}\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right) = \frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right)\frac{1}{\theta f(\theta)}$$
$$\times \left[\int_{\theta}^{\bar{\theta}} \frac{1}{u'\left(c\left(\hat{\theta}\right)\right)} f\left(\hat{\theta}\right) d\hat{\theta} - \frac{1-H(\theta)}{\lambda}\right] \left[\frac{y(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)} + 1\right].$$

Next, note that the term on the left hand side of this equality captures the intratemporal wedge between consumption and leisure in a decentralized version of this economy. The reason is, if we let $T(y) = a + \tau y$ be the tax rate an individual of type θ faces in a decentralized market, his first order condition for consumption-leisure choice would be:

$$\frac{\frac{1}{\theta}v'\left(\frac{y}{\theta}\right)}{u'(c)} = 1 - \tau.$$

Therefore, if we let

$$1 - \tau(\theta) \equiv \frac{v'\left(\frac{y(\theta)}{\theta}\right)}{\theta u'(c(\theta))},$$

we can write:

$$\frac{\tau(\theta)}{(1-\tau(\theta))} = \frac{[1-F(\theta)]}{\theta f(\theta)}\left[1 + \frac{y(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)}\right]$$
$$\times \left[\int_{\theta}^{\bar{\theta}} \frac{u'(c(\theta))}{u'\left(c\left(\hat{\theta}\right)\right)}\frac{f\left(\hat{\theta}\right)d\hat{\theta}}{[1-F(\theta)]} - \frac{u'(c(\theta))}{\lambda}\frac{1-H(\theta)}{[1-F(\theta)]}\right]. \quad (9)$$

The left hand side of this equality is increasing in τ. Therefore, it gives us a means for studying the behavior of optimal marginal tax rates across the types, and how they depend on the parameters of the problem. Every change that leads to an increase in the right hand side of this equality implies an increase in the distortion at the margin.

The first term on the right hand side of this equality is called the *hazard rate*; it is a function of the tail of the distribution of types. The second term depends on the elasticity of labor supply. To see why, suppose $v(l) = l^{\epsilon}$, for some $\epsilon > 1$. Then, we have:

$$\frac{y(\theta)}{\theta} \frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)} = \epsilon - 1,$$

which is the inverse of Frisch elasticity of labor supply.

Remark 13 *Frisch elasticity of labor supply is the elasticity of labor supply with respect to wage, while the marginal utility of wealth is kept constant. In other words, it captures the substitution effect of a marginal change in wage. In the case of our economy, given this form of disutility of labor, the problem of an individual becomes:*

$$\max_{c,l} \; u(wl) - l^{\epsilon}.$$

The first order conditions for this problem imply:

$$l^{\epsilon-1} = \frac{wu'(wl)}{\epsilon}.$$

Therefore:

$$\log(l) = \log\left(\frac{u'(wl)}{\epsilon}\right)^{\frac{1}{\epsilon-1}} + \frac{1}{(\epsilon-1)} \log(w).$$

If the marginal utility of wealth, $u'(wl) = u'(c)$ is kept constant, the elasticity of labor with respect to wage rate is $1/(\epsilon - 1)$.

Then, we can write the above equality as:

$$\frac{\tau(\theta)}{(1-\tau(\theta))} = \frac{[1-F(\theta)]}{\theta f(\theta)} \epsilon \left[\int_{\theta}^{\bar{\theta}} \frac{u'(c(\theta))}{u'\left(c\left(\hat{\theta}\right)\right)} \frac{f\left(\hat{\theta}\right) d\hat{\theta}}{[1-F(\theta)]} - \frac{u'(c(\theta))}{\lambda} \frac{1-H(\theta)}{[1-F(\theta)]} \right].$$

However, we cannot say, by looking at this equality, that a change in the tail of the type distribution, or a change in the Frisch elasticity, ϵ, would change the tax rate in a certain way. The reason is that all the parameters on the right hand side of this equality are inter-connected; a change in F or ϵ would cause everything to change.

However, we are still able to study special cases of interest.

2.1 Risk-Neutral Agents

Suppose utility of consumption is given by $u(c) = c$. Then, in the case of $v(l) = l^{\epsilon}$, (9) becomes:

$$\frac{\tau(\theta)}{(1-\tau(\theta))} = \frac{1}{\theta f(\theta)} \epsilon \left[1 - F(\theta) - \frac{1-H(\theta)}{\lambda}\right].$$

An increase in ϵ would lead to an increase in τ for all types. On the other hand, an increase in H (in a first order stochastic dominance sense) would make the tax rate on higher types to increase.

2.2 Type Distribution with Pareto Tail

Suppose the distribution of types has the form $1 - F(\theta) = C\theta^{-a}$. Then:

$$\frac{1 - F(\theta)}{\theta f(\theta)} = \frac{C\theta^{-a}}{\theta a C\theta^{-a-1}} = \frac{1}{a}.$$

If $y(\theta) = \bar{l}\theta$, then the income distribution would be a good measure of the distribution of types in the data.

This section remains to be completed!

3 Sufficiency of Local Incentive Constraints

(This Appendix was written by Anderson Schneider.)

In this section, we show that only local incentive constraints matter, when solving the planning problem in the presence of private information.

Utility take the form $U(c, l) = u(c) - v(l)$, with $u(\cdot)$ being strictly concave, and $v(\cdot)$ strictly convex.

Lemma 14 *Under Assumption 3, for any $\theta_j > \theta_i$, we have:*

1. *if $u(c) - v(y/\theta_i) \geq u(\hat{c}) - v(\hat{y}/\theta_i)$, $c \geq \hat{c}$, and $y \geq \hat{y}$, then*

$$u(c) - v(y/\theta_j) \geq u(\hat{c}) - v(\hat{y}/\theta_j);$$

2. *if $u(c) - v(y/\theta_j) \geq u(\hat{c}) - v(\hat{y}/\theta_j)$, $c \leq \hat{c}$, and $y \leq \hat{y}$, then*

$$u(c) - v(y/\theta_i) \geq u(\hat{c}) - v(\hat{y}/\theta_i).$$

Proof. We will prove only Part 1; Part 2 follows a similar argument (one with reversed signs).

Suppose, for the sake of contradiction, that this was not the case. Then, there would exist $(c, y) \geq (\hat{c}, \hat{y})$, and $\theta_j > \theta_i$, such that

$$u(c) - v\left(\frac{y}{\theta_i}\right) \geq u(\hat{c}) - v\left(\frac{\hat{y}}{\theta_i}\right). \tag{10}$$

But, then,

$$\begin{aligned} u(c) - v\left(\frac{y}{\theta_j}\right) &< u(\hat{c}) - v\left(\frac{\hat{y}}{\theta_j}\right), \\ \therefore u(c) - u(\hat{c}) &< v\left(\frac{y}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_j}\right). \end{aligned} \tag{11}$$

Now, define the function

$$F(x, w_i, w_j) \equiv v\left(\frac{x}{w_j}\right) - v\left(\frac{x}{w_i}\right),$$

and notice that

$$\frac{\partial F(x, w_i, w_j)}{\partial x} = v'\left(\frac{x}{w_j}\right)\frac{1}{w_j} - v'\left(\frac{x}{w_i}\right)\frac{1}{w_i} < 0,$$

where the last inequality follows from the convexity of $v(\cdot)$. Then, since $F(x, w_i, w_j)$ is decreasing in x, it follows that

$$F(x, \theta_i, \theta_j) \leq F(\hat{x}, \theta_i, \theta_j).$$

Therefore

$$\begin{aligned}
\therefore v\left(\frac{y}{\theta_j}\right) - v\left(\frac{y}{\theta_i}\right) &\leq v\left(\frac{\hat{y}}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_i}\right), \\
\therefore v\left(\frac{\hat{y}}{\theta_i}\right) - v\left(\frac{y}{\theta_i}\right) &\leq v\left(\frac{\hat{y}}{\theta_j}\right) - v\left(\frac{y}{\theta_j}\right), \\
\therefore v\left(\frac{y}{\theta_i}\right) - v\left(\frac{\hat{y}}{\theta_i}\right) &\geq v\left(\frac{y}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_j}\right).
\end{aligned}$$

Using the last inequality, together with (10), we get:

$$u(c) - u(\hat{c}) \geq v\left(\frac{y}{\theta_i}\right) - v\left(\frac{\hat{y}}{\theta_i}\right) \geq v\left(\frac{y}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_j}\right)$$

which contradicts (11). ■

It turns out that a strict version for Lemma 14 also holds.

Lemma 15 *Under Assumption 3, for any $\theta_j > \theta_i$, we have:*

1. *if $u(c) - v(y/\theta_i) > u(\hat{c}) - v(\hat{y}/\theta_i)$, $c \geq \hat{c}$, and $y \geq \hat{y}$, then*

$$u(c) - v(y/\theta_j) > u(\hat{c}) - v(\hat{y}/\theta_j);$$

2. *if $u(c) - v(y/\theta_j) > u(\hat{c}) - v(\hat{y}/\theta_j)$, $c \leq \hat{c}$, and $y \leq \hat{y}$, then*

$$u(c) - v(y/\theta_i) > u(\hat{c}) - v(\hat{y}/\theta_i).$$

Proof. We will prove only Part 1; Part 2 follows a similar argument (one with reversed signs).

Suppose, for the sake of contradiction, that this was not the case. Then, there would exist $(c, y) \geq (\hat{c}, \hat{y})$, and $\theta_j > \theta_i$, such that

$$u(c) - v\left(\frac{y}{\theta_i}\right) > u(\hat{c}) - v\left(\frac{\hat{y}}{\theta_i}\right). \tag{12}$$

But, then,

$$\begin{aligned} u(c) - v\left(\frac{y}{\theta_j}\right) &\leq u(\hat{c}) - v\left(\frac{\hat{y}}{\theta_j}\right), \\ \therefore u(c) - u(\hat{c}) &\leq v\left(\frac{y}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_j}\right). \end{aligned} \tag{13}$$

Now, define the function

$$F(x, w_i, w_j) \equiv v\left(\frac{x}{w_j}\right) - v\left(\frac{x}{w_i}\right),$$

and notice that

$$\frac{\partial F(x, w_i, w_j)}{\partial x} = v'\left(\frac{x}{w_j}\right)\frac{1}{w_j} - v'\left(\frac{x}{w_i}\right)\frac{1}{w_i} < 0,$$

where the last inequality follows from the convexity of $v(\cdot)$. Then, since $F(x, w_i, w_j)$ is decreasing in x, it follows that

$$F(x, \theta_i, \theta_j) \leq F(\hat{x}, \theta_i, \theta_j).$$

Therefore

$$\begin{aligned} v\left(\frac{y}{\theta_j}\right) - v\left(\frac{y}{\theta_i}\right) &\leq v\left(\frac{\hat{y}}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_i}\right), \\ \therefore v\left(\frac{\hat{y}}{\theta_i}\right) - v\left(\frac{y}{\theta_i}\right) &\leq v\left(\frac{\hat{y}}{\theta_j}\right) - v\left(\frac{y}{\theta_j}\right), \\ \therefore v\left(\frac{y}{\theta_i}\right) - v\left(\frac{\hat{y}}{\theta_i}\right) &\geq v\left(\frac{y}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_j}\right). \end{aligned}$$

Using the last inequality, together with (12), we get:

$$u(c) - u(\hat{c}) > v\left(\frac{y}{\theta_i}\right) - v\left(\frac{\hat{y}}{\theta_i}\right) > v\left(\frac{y}{\theta_j}\right) - v\left(\frac{\hat{y}}{\theta_j}\right)$$

which contradicts (13). ■

Lemma 16 *Under Assumption 3,*

$$u\left(c\left(\theta_i\right)\right)-v\left(\frac{y\left(\theta_i\right)}{\theta_i}\right)\geq u\left(c\left(\theta_j\right)\right)-v\left(\frac{y\left(\theta_j\right)}{\theta_i}\right),\ \textit{for } j=i-1,\ i+1,$$

implies

$$u\left(c\left(\theta_i\right)\right)-v\left(\frac{y\left(\theta_i\right)}{\theta_i}\right)\geq u\left(c\left(\theta_j\right)\right)-v\left(\frac{y\left(\theta_j\right)}{\theta_i}\right),\quad \forall j.$$

Proof. Suppose, for the sake of contradiction, that this was not the case, and, without loss of generality, assume that is not true for $j=i+2$. Then, we would have:

$$u\left(c\left(\theta_i\right)\right)-v\left(\frac{y\left(\theta_i\right)}{\theta_i}\right)\geq u\left(c\left(\theta_{i+1}\right)\right)-v\left(\frac{y\left(\theta_{i+1}\right)}{\theta_i}\right),$$

and

$$u\left(c\left(\theta_{i+2}\right)\right)-v\left(\frac{y\left(\theta_{i+2}\right)}{\theta_i}\right)> u\left(c\left(\theta_i\right)\right)-v\left(\frac{y\left(\theta_i\right)}{\theta_i}\right).$$

These inequalities imply that

$$u\left(c\left(\theta_{i+2}\right)\right)-v\left(\frac{y\left(\theta_{i+2}\right)}{\theta_i}\right)> u\left(c\left(\theta_{i+1}\right)\right)-v\left(\frac{y\left(\theta_{i+1}\right)}{\theta_i}\right). \tag{14}$$

By Lemma 15, we know that $\left(c\left(\theta_{i+2}\right),y\left(\theta_{i+2}\right)\right)>\left(c\left(\theta_{i+1}\right),y\left(\theta_{i+1}\right)\right)$. Since $\theta_{i+1}>\theta_i$, we can apply Part 1 of Lemma 15 to Equation (14) to get:

$$u\left(c\left(\theta_{i+2}\right)\right)-v\left(\frac{y\left(\theta_{i+2}\right)}{\theta_{i+1}}\right)> u\left(c\left(\theta_{i+1}\right)\right)-v\left(\frac{y\left(\theta_{i+1}\right)}{\theta_{i+1}}\right),$$

which violates the incentive constraint for type θ_{i+1}.

Now, suppose that was not true for $j=i-2$. Then, as before, we would have:

$$u\left(c\left(\theta_i\right)\right)-v\left(\frac{y\left(\theta_i\right)}{\theta_i}\right)\geq u\left(c\left(\theta_{i-1}\right)\right)-v\left(\frac{y\left(\theta_{i-1}\right)}{\theta_i}\right),$$

and

$$u\left(c\left(\theta_{i-2}\right)\right)-v\left(\frac{y\left(\theta_{i-2}\right)}{\theta_i}\right)> u\left(c\left(\theta_i\right)\right)-v\left(\frac{y\left(\theta_i\right)}{\theta_i}\right).$$

These inequalities imply that

$$u\left(c\left(\theta_{i-2}\right)\right)-v\left(\frac{y\left(\theta_{i-2}\right)}{\theta_i}\right)> u\left(c\left(\theta_{i-1}\right)\right)-v\left(\frac{y\left(\theta_{i-1}\right)}{\theta_i}\right). \tag{15}$$

Since by Lemma 15, $\left(c\left(\theta_{i-2}\right),y\left(\theta_{i-2}\right)\right)<\left(c\left(\theta_{i-1}\right),y\left(\theta_{i-1}\right)\right)$, and because $\theta_{i-1}<\theta_i$, using Part 2 of this lemma, Equation (15) implies:

$$u\left(c\left(\theta_{i-2}\right)\right)-v\left(\frac{y\left(\theta_{i-2}\right)}{\theta_{i-1}}\right)> u\left(c\left(\theta_{i-1}\right)\right)-v\left(\frac{y\left(\theta_{i-1}\right)}{\theta_{i-1}}\right),$$

which violates the incentive constraint for type θ_{i-1}. ■