3 Optimal income taxation

3.1 Introduction

It is interesting to extend the optimal tax model to deal with income inequality and progressive taxes on (labour) income. In fact, we have already studied a *negative income tax* of the form $T(z) = -G + tz$ where $z = wh$ is labour income. Recall that the optimal commodity tax with heterogeneous households and a uniform lump-sum tax/transfer can be interpreted in this way. There we showed the optimal linear tax rate satisfied (in the case of two goods at least)

$$\frac{t^* \bar{z}_t}{\bar{z}} = \text{cov}(b^h, \frac{z^h}{\bar{z}})$$

We can do comparative statics on social preferences, inequality, and behavioural responses with this model.

3.2 Non-linear income taxation: The two-type case

Here we offer a less technical version of the optimal non-linear income tax, based on the two-class economy studied by Stiglitz (1982).

An individual taxpayer receives a pre-tax wage rate $s \in \{s_l, s_h\}$ with $s_l < s_h$. A fraction p_i of individuals of type $i = l, h$. Each taxpayer works y hours to earn pre-tax income $z = sy$. Preferences over consumption x and labour hours y given by a quasi-concave function $V(x, y)$.

The social planner chooses a tax function $T(z)$, yielding after-tax income $x = z - T(z)$ for a taxpayer with pre-tax income z. Note that facing this tax schedule an agent chooses (x_i, z_i) to $\max V(z - T(z), z/s)$ or $V_y/V_x = s(1 - T')$.

The tax payments for each type $i = l, h$ are chosen to maximize a social welfare function $W(u_l, u_h)$. If W is concave, therefore, the goal is in effect to redistribute income from high-wage to low-wage in an efficient manner.

To understand the problem, consider first the case in which wage rates s_i of each taxpayer are observable to the planner. Then lump-sum redistribution between the two types is feasible: if the planner wishes to make type l better off, then each type h pays a lump-sum tax in amount T, and each type l receives a lump-sum transfer $S = p_h T / p_l$.

The problem: potential market wage rates s_i are unobserved. Direct inferences of whether income z is low because of y or s are not possible. (But what if y_i could be observed?) Heuristically, then, we wish to design a tax function T to reduce incentive for type h individuals to choose low y and escape taxation.

The problem can most conveniently be represented as one of choosing allocations $\alpha_i = (x_i, z_i)$ for each type to maximize social welfare subject to the government budget constraint and to the incentive (self-selection) constraint for type h:

$$V(x_h, z_h/s_h) \geq V(x_l, z_l/s_h)$$
$$V(x_l, z_l/s_l) \geq V(x_h, z_h/s_l)$$

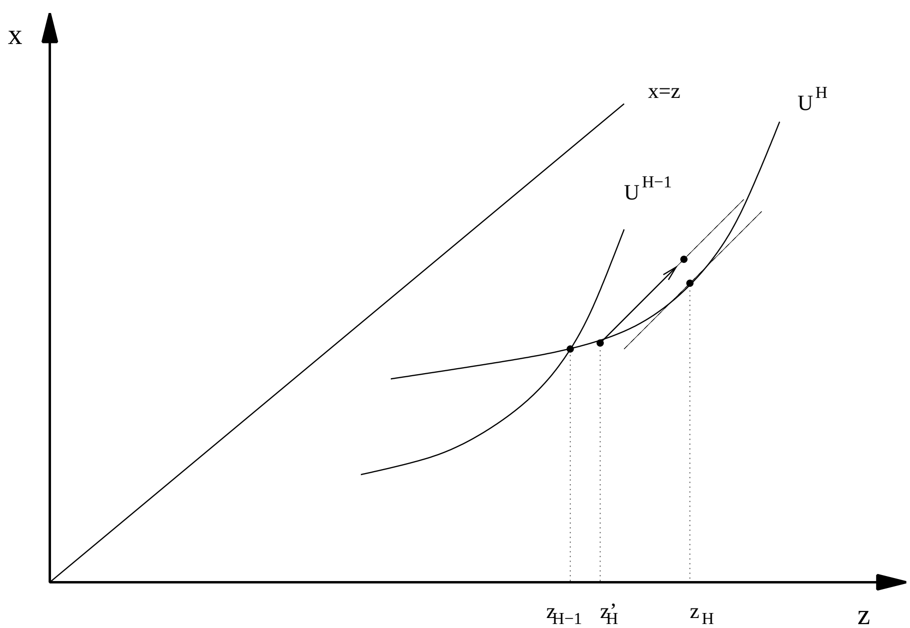

Figure 3.1: No distortion at the top.

A nice feature of this problem is that, under weak conditions on V, the indifference curves in (x,z)-space for the h-type are always flatter than those of the l-type. This reflects the fact that any (x,z) combination involves less labour for the h-type, and so marginal increases in z should be less costly for h too. Under this *single-crossing* or *agent monotonicity* property, we know:

- type h must be offered higher pre-tax income than l, i.e. $z_h \geq z_l$ at all incentive compatible allocations; and
- if the incentive constraint for h is satisfied then the IC for l will be satisfied as a strict inequality, and vice versa.

The proof of these statements is in the pictures.

A sufficient condition for SCP here is that consumption be a normal good. To see this, note the MRS is

$$\sigma = -\frac{V_z}{V_x} = -\frac{V_y(x, z/s)}{V_x(x, z/s)s}$$

So σ is decreasing in s if $-V_y/V_x$ is decreasing in s, which holds if the same ratio is increasing in $y = z/s$. But this is just the standard condition for an increase in lump-sum income to increase consumption: normality of x.

In a typical solution to this problem (i.e. when the planner wishes to redistribute from h to l, and a first-best allocation is not implementable), it is the first constraint will be binding at the optimum, so the second will be slack. We then have the following result for the marginal tax rates on labour income $T'(z_i)$ faced by each type at the optimum. Formally,

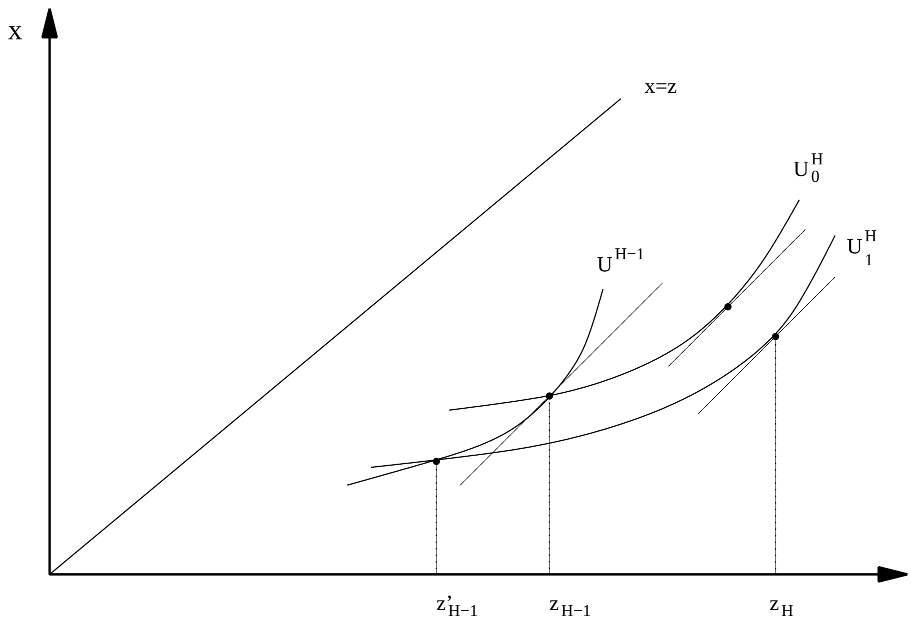

Figure 3.2: Direction of distortion at binding incentive constraints.

Definition 1 *A* typical solution *to the optimal tax problem is one at which the planner would strictly prefer to make a small lump sum transfer from h to l, if incentive constraints could be ignored.*

Proposition 1 *In a typical solution to the optimal tax problem, $T'(z_h) = 0$ and $T'(z_l) > 0$.*

We can provide a simple graphical proof for the proposition in (x, z)-space. To illustrate the first result, suppose that a bundle α_l has been chosen for l. We claim α_h should be at a point tangent to a 45-degree line, so that $T'_h = 0$. By assumption, IC for h to l is binding. Suppose $T'_h \neq 0$. (See Figure 3.1.) So we can move along a 45-degree line keeping revenue constant and increasing welfare for h. Since IC is slack for l, this is incentive compatible and Pareto-improving: a contradiction.

For the second result, now consider solution for l. Suppose $T'_l \leq 0$. (See Figure 3.2.) Consider a movement along indifference curve for l toward positive marginal tax rate. Along this path, transfer to l is non-increasing, so the change is feasible, and u_l is constant. But the change in α_l causes an increase in the incentive-compatible revenue extracted from h. Since by assumption the planner prefers to do more redistribution if feasible, the original tax schedule was not an optimum: a contradiction.

Note the result of "no distortion at the top" is rather different than what is observed in actual tax schedules. The result that marginal tax rates should be positive at lower income levels is as expected. Note however that the reason for a high marginal tax rate on l is to extract greater revenue from h, not l.

3.2.1 Extension to many types

With a finite number of types greater than two, the analysis is qualitatively similar. Each type is assigned a bundle α, and (if redistributive preferences are strong) bundles are connected by a change of downward-binding incentives constraints. Since there is no binding constraint at the bundle offered to the highest-wage type, this bundle is undistorted, i.e. it involves a zero marginal tax rate.

With many types, we must take seriously the possibility that many types are "bunched" at a single α. The analysis of when this occurs is complicated. Intuitively, it may not be worthwhile to separate two types if the desire to redistribute *between these two types* is small, because this will require distorting the bundles offered to all types with higher wages. So, bunching is likely to occur for very low wages. And bunching is likely to occur at wages that are earned by a small fraction of the population.

3.3 Continuous wage distributions

We now consider a more general formal analysis of the optimal income tax problem. This is adapted and simplified from Diamond (1998).

Let s denote a particular individual's pre-tax wage rate, $x(s)$ her consumption and $y(s)$ her labour supply. The utility function is

$$V(x,y) = x - \frac{y^{1+\epsilon}}{1+\epsilon} \tag{3.1}$$

Note that this utility function is quasi-linear in consumption: there are no income effects on labour supply, which simplifies the analysis.

Suppose the agent were faced with a non-linear tax schedule $T(z)$ where $z = sy$. Her problem would be to $\max_y sy - T(sy) - y^{1+\epsilon}/(1+\epsilon)$, yielding first-order condition

$$s(1-T') = y^{\epsilon} \tag{3.2}$$

That is, $1/\epsilon$ is the elasticity of labour supply with respect to the after-tax wage rate.

The planner observes each agent's labour income $z = sy$ and knows the distribution of wage rates in society, which varies on the support $[s_0, s_1]$ and has c.d.f. $F(s)$ and density $f(s)$. However, the planner cannot observe any individual's wage rate not her labour supply y (is this reasonable?). The goal is therefore to design a tax function $T(sy)$ to redistribute income (in a way we make precise shortly).

Incentive compatibility: The first-order approach. This can be represented as a problem in which each taxpayer reports a wage rate s and is assigned a labour supply $y(s)$ and consumption level $x(s)$. (That is, we will solve for the *optimal direct mechanism*.) The question is, will each taxpayer report her wage s truthfully? Let

$$U(s',s) = x(s') - \frac{1}{1+\epsilon}\left(\frac{s'y(s')}{s}\right)^{1+\epsilon}$$

denote the payoff wage s reporting any wage s', and let $u(s) = U(s,s)$ denote the payoff to truthful reporting. If truth-telling is optimal, then for all s

$$\begin{aligned} U_1(s,s) &= 0 \\ U_{11}(s,s) &= -U_{12}(s,s) \le 0 \end{aligned}$$

where the equality in the second line follows from differentiating the first. To see the implications of the truth-telling conditions, first differentiate to get

$$U_2(s', s) = (z(s'))^{1+\epsilon} s^{-(\epsilon+2)}$$

Applying the FOC above implies (envelope theorem):

$$u'(s) \equiv \frac{du(s)}{ds} \equiv U_2(s,s) = y(s)^{1+\epsilon} s^{-1} \tag{3.3}$$

Applying the SOC implies

$$U_{11}(s,s) = -U_{21}(s,s) \leq 0 \quad \Longleftrightarrow \quad z'(s') \geq 0 \tag{3.4}$$

Equations (3.3) and (3.4) are necessary conditions for truth-telling, describing how utility must rise with the wage rate in the equilibrium. A more detailed analysis would show that these condition is also sufficient for truth-telling – in the sense that any allocation satisfying them can be made incentive compatible by suitable choice of $x(s)$.

3.3.1 The optimal tax system

The problem is to maximize an additive, concave social welfare function subject to the government budget constraint and the incentive constraint. Thus:

$$\begin{aligned} \max \int_{s_0}^{s_1} \frac{u(s)^{1+\theta}}{1+\theta} f(s) ds \\ \text{subject to } \int_{s_0}^{s_1} [x(s) - sy(s)] f(s) ds + R_0 \leq 0 \\ u'(s) = y^{1+\epsilon} s^{-1} \end{aligned}$$

Some notes on this problem:

1. We should assume $\theta < 0$ for a concave SWF.
2. Why do we write the GBC as this material balance condition? Note that $x + T = sy$ so $\int T f(s) ds = \int [sy - x] f(s) ds$.
3. We are ignoring the non-negativity constraint $z' \geq 0$: we are assuming the first-order approach is valid. Later we can check whether our solution actually satisfies the constraint.

The last step is to eliminate consumption $x(s)$ from the problem. The utility function implies $x = u + y^{1+\epsilon}/(1+\epsilon)$ so the GBC is now

$$\int_{s_0}^{s_1} [u(s) + y(s)^(1+\epsilon)/(1+\epsilon) - sy(s)] f(s) ds + R_0 \leq 0$$

To solve the problem, form the Hamiltonian

$$H = \frac{u^{1+\theta}}{1+\theta} f + \mu \left(sy - \frac{y^{1+\epsilon}}{1+\epsilon} - u \right) f + \Lambda u'$$

where Λ is the co-state variable for the state u, and y is the control variable. The first-order necessary conditions are

$$H_u = (u^\theta - \mu)f = -\Lambda' \tag{3.5}$$

$$H_y = \mu(s - y^\epsilon)f + \Lambda(1+\epsilon)y^\epsilon s^{-1} = 0 \tag{3.6}$$

Observe that $\Lambda(s_1) = 0$ since the endpoint utility is a free variable, so we can integrate (3.5) to get

$$\Lambda(s) = -\int_s^{s_1} \Lambda'(w)\,dw = -\int_s^{s_1} (\mu - u^\theta) f\,dw$$

To interpret the expression, recall (3.2) tells us $s(1-T') = y^\epsilon$ or $s - y^\epsilon = sT'$. Substitute into (3.6) to get[1]

$$\frac{T'}{1-T'} = -\frac{(1+\epsilon)\Lambda(s)}{\mu s f(s)} = (1+\epsilon)\frac{1-F(s)}{sf(s)}\frac{\int_s^{s_1}(\mu - u^\theta)f(w)\,dw}{\mu(1-F(s))} \tag{3.7}$$

The left-hand side of (3.7) is an increasing function of the marginal tax rate faced by an individual with wage s. The three terms on the right-hand side capture the various tradeoffs involved in increasing the marginal tax rate at a given point in the wage distribution.

The first term on the right-hand side is a decreasing function of the labour supply elasticity. Thus taxes should be lower where the distortionary consequences are larger. Increasing the marginal tax rate at a given point in the distribution induces an efficiency cost that is proportional to the proportion of the population that are directly affected by it $f(s)$ and the value of their labour for income tax revenues s, which explains the denominator of the second term. On the other hand, increasing the marginal tax rate at a given point in the distribution imposes a higher average tax rate on all higher wage earners and so induces greater redistribution, which explains the numerator of the third term. (More formally, it measures the value of redistribution away from those above wage s.) Thus the marginal tax rate should be increasing in the ratio of these two expressions. Thus the expression says that tax rates should higher when the elasticity of labour supply is smaller and at wage rates where there are relatively few people but there are many people with higher wage rates. All this is consistent with our two-type graphical analysis.

3.4 Implications

The shape of the marginal tax rate schedule. Interpreting (3.7) is hard in general, and learning about the shape and level of the MTR schedule requires considering special cases and often numerical methods. Diamond (1998) argues that, in plausible cases, the MTR schedule is "U"-shaped—the marginal tax rate is initially decreasing in earned income (at least above some earnings level) and then increasing for higher earnings.

[1] Intermediate step for above: since

$$1 - T' = y^\epsilon s^{-1}$$
$$T' = (s - y^\epsilon)s^{-1}$$

we have

$$\mu s T' f + (1+\epsilon)(1-T')\Lambda = 0$$

I sketch the argument here; more details below. The third term on the right-hand side of (3.7) is increasing in s; and the first term is constant.

The second term may be increasing or decreasing in general. Suppose that the skill distribution fits the Pareto distribution above some wage level s^*, i.e. $1 - F(s) = (\beta/s)^\alpha$. Then the second term can be shown to be constant above s^*, and T' is increasing in s above s^*. Likewise, Diamond argues that T' must be decreasing in s for skill levels below the mode, which gives the "U"-shape (perhaps with some wobbles) overall. Thus the high marginal tax rates we observe in western countries on both high earners and net transfer recipients may in principle be optimal.

To understand more about the shape of the MTR schedule, consider the third term. Notice this term is proportional to the average value of $\mu - u^\theta$ above s:

$$E(\mu - u(w)^\theta | s > w) = \frac{\int_s^{s_1} (\mu - u^\theta) f \, dw}{1 - F(s)}$$

However, incentive compatibility requires u be increasing in s, and $\theta < 0$ implies $\mu - u^\theta$ is increasing in s. Hence the truncated mean is also increasing in s.

On the other hand, the second term may be increasing or decreasing in s, depending on the underlying wage distribution. One important (and easy) case is when the wage distribution follows a Pareto distribution for wages above some level β up to the maximal wage s_1. Empirical studies have shown that actual income distributions above the mode are well fitted by the Pareto family. The c.d.f. for the Pareto is given by

$$F(s) = 1 - (\beta/s)^\alpha \qquad (\alpha > 0)$$

for which the second term is a constant (since $f = \alpha\beta^\alpha s^{-\alpha-1}$ implies $(1-F)/sf = 1/\alpha$).

Of course, very different tax rate schedules are possible, but it is interesting that marginal progressivity emerges under such reasonable conditions. Diamond extends the example to show the possibility of a u-shape of marginal tax rates. This relies on the fact that the numerator of the third term $N = \int_s^{s_1} (\mu - u^\theta) f \, dw$ has a unique maximum for an interior wage rate, say at a wage s_c defined by

$$u(s_c)^\theta = \mu$$

Since the denominator of the second term is increasing below the mode of s (say s_m), if $s_c < s_m$ then marginal tax rates must be falling in $[s_c, s_m]$. Again, the example is rather special, but it is interesting to see that a system with highest disincentives for the richest and poorest alike can be rationalized by the model.

Marginal tax rates on high earners. Inspecting (3.7), we see that $T'(s_1) = 0$, replicating our earlier result of "no distortion at the top". But the preceding argument for T' to be increasing above s^* holds in any open interval below s_1, showing that the limit result is very special indeed.

How high can marginal tax rates on high earners get? To answer this, we can let s_1 grow large and take the limit of tax rates $T'(s_1 - \delta)$ for fixed $\delta > 0$. Since $u(s)$ is likely unbounded, the limit of the third term in (3.7) approaches unity, and

$$\frac{T'}{1 - T'} \to \frac{1 + \epsilon}{\alpha}$$

For ϵ, values in $[1, 10]$ seem reasonable (a right-winger is on the left boundary, and conversely). For α, values in $[1, 2]$ are most reasonable, since it is only in this range that the first two central moments of the distribution are defined. This gives values for the limiting expression in $[1, 11]$, and asymptotic tax rates between 50 and 92 per cent.

3.5 Applications and extensions

3.5.1 The role of commodity taxes

Given an optimal non-linear income, should the planner employ (linear or non-linear) commodity taxes as well? Since we have shown that a proportional tax on labour income is equivalent to a uniform linear tax on all net commodity demands, the question is really whether differentiated commodity taxation is desirable, given optimal income taxation. Let $\mathbf{x}$ be an n-vector of individual commodity demands, y be labour supply, and $V(\mathbf{x}, y)$ be the utility function (not necessarily quasi-linear as in the Diamond example above). Atkinson and Stiglitz (1976) showed that if the utility function can be written in the weakly separable form $V(\mathbf{x}, y) = G(H(\mathbf{x}), y)$, then there is no gain to differentiated commodity taxation. The intuition is that weak separable in labour/leisure implies marginal rates of substitution with respect to commodity demands are independent of labour. Thus commodity demands contain no information about an individual's labour supply and hence no information about the unobservable skill parameter. It follows distorting commodity demands can do nothing to relax the incentive constraints in the optimal tax problem, and so should be avoided.

3.5.2 Optimal family taxation

What is the appropriate unit of taxation, the individual or the family? There are equity and efficiency aspects.

Horizontal equity. It seems plausible that welfare is a function of the sum of earnings within the family, possibly with some adjustment for family size (adult equivalence scales). This creates conflicts between vertical and horizontal equity in the tax system: suppose we wished to design a tax system $T_s(y)$ for singles and $T_m(y_1, y_2)$ for couples with the properties of:

1. equal treatment of one-earner and two-earner couples: $T_m(y_1, y_2) = T^*(y_1 + y_2)$ for all (y_1, y_2);
2. increasing marginal tax rates for all (y_1, y_2); and
3. marriage neutrality: $T_s(y_1) + T_s(y_2) = T_m(y_1, y_2)$.

It is easy to see that no such tax system exists. That is $T_s(y_1) + T_s(y_2) = T^*(y_1 + y_2)$ can only be satisfied when taxes are affine: $T_s(y) = \alpha + \beta y$. Intuitively, a tax based on family income alone (whether average or total income or whatever) must subsidize or penalize marriage if there is any marginal progressivity. More on this in Berliant and Rothstein (NTJ, 2003).

Efficiency. On efficiency grounds, it is often argued (e.g. Boskin and Sheshinski, 1983) that the compensated labour supply elasticities of married women are much higher than those of married men (chiefly because participation is elastic for women but not men). Since married women (still) earn less on average than the men they marry, there is a prima facie efficiency case for individual rather than family taxation, i.e. abandoning principle 1.

On the other hand, Piggott and Whalley (1996) argue that individual-based progressive taxation distorts the relative opportunity costs of spouses in supplying home production services rather than market labour (if men face higher tax rates then they do too much housework!). Fixing this distortion requires family based taxation.

In an interesting paper, Kleven, Kreiner and Saez (2007) study the optimal *non-linear tax* on families. They show that it is generally optimal for the tax rate on secondary earners to depend on the income of the primary earner. This doesn't really contradict Boskin–Sheshinski, and it is a bit obvious: the planner would like to redistribute to one-earner from two-earner families at the same total income, and will tolerate some efficiency loss. More interesting, they argue that the optimal MTR premium for married earners is *decreasing* in family income. This is the opposite of what we see in progressive family-based tax systems (except for benefit clawbacks); the intuition is largely the same as above.

3.6 General equilibrium effects

In the above analysis, we assumed that wage rates were fixed and equal to "skills," independent of the level of taxation. Stiglitz (1982) considered the two-type model, in which output was some function $G(y_l, y_h)$ of the labour effort of the two types. He showed that, for *any* constant-returns production not equal to the linear function $G(y_l, y_h) = s_l y_l + s_h y_h$, it is optimal to impose a negative marginal tax rate on the high-ability type. Naito (1999) and others have extended the analysis to show that a variety of other distortions are desirable, if they increase pre-tax redistribution from high to low types.

3.7 Low-income support and work incentives

Suppose that agents differ in their attitudes to leisure, as well as their abilities. Then low observed pre-tax income might reflect low skill or high taste for leisure, and the planner might in principle assign very different welfare weights to consumption by the two types of agent. Analyzing the optimal tax schedule in this context is very difficult, since with multidimensional types, the single crossing property cannot be used to guarantee that adjacent downward incentive constraints, and only these incentive constraints, are binding.

Some progress on the problem was made by Saez (QJE, 2002).[2] He characterizes the optimal tax system when government may assign greater welfare weight to those working than those unemployed, and shows it may involve negative marginal tax rates on low earners—an "earned income tax credit" that subsidizes work by transfer recipients. A version of this argument is in Kanbur, Keen, and Tuomala (JPubE, 1994), who simply assume that government minimizes poverty (so does not care at all about leisure).

The same kind of arguments can be made to support workfare – again, causing transfer recipients to work more than is really productive. Thus Besley and Coate (REStud, 1989) look at poverty reduction and Cuff (CJE, 2001) looks at welfare optimums.

The last related issue is the desirability of a minimum wage: why employ this means of redistribution when it is productively inefficient? Boadway and Cuff (2001) argue that it is part of efficient redistribution because it targets low *wage* agents rather than low *income* people, thereby getting around the inefficiency built into the Mirrlees model. Saez (2007) extends this argument to explain why the planner would choose to ignore wage information for osme if it were available.

Less explored is the role of in-kind redistribution. Traditionally economists denigrate programs like housing subsidies, free health care and so on as an inefficient means of redistribution; the traditional view has been that these are *merit goods*. Blackorby and Donaldson (AER, 1988) analyze how in-kind redistribution can be second-best efficient because it relaxes self-selection constraints.

[2] Emmanuel Saez, 2002, "Optimal income transfer programs: Intensive versus extensive labor supply responses," *Quarterly Journal of Economics* 117, 1039–1073.

But there is little work that attempts to estimate the gain in redistribution (labour) efficiency against the loss in consumption efficiency, and so on.

Appendix

First we derive the optimal tax rate expression using the method of Lagrange rather than the optimal control method.

Since $x = u + y^{1+\epsilon}/(1+\epsilon)$, the material balance condition becomes

$$\int_{s_0}^{s_1} [sy(s) - \frac{y(s)^{1+\epsilon}}{1+\epsilon} - u(s)] f(s) ds \geq R_0 \tag{3.8}$$

If the incentive constraint $u' = y^{1+\epsilon}/n$ holds, then the utility obtained at the optimum by any taxpayer can be written as

$$u(s) = u(s_0) + \int_{s_0}^{s} u'(w) dw = u(s_0) + \int_{s_0}^{s} y(w)^{1+\epsilon} w^{-1} dw \tag{3.9}$$

To solve the problem, let μ denote the multiplier on the material balance condition (3.8), and let $\lambda(n)$ denote the multiplier on the utility constraint (3.9). Then form the Lagrangian

$$\begin{aligned}\mathscr{L} = \int_{s_0}^{s_1} \frac{u(s)^{1+\theta}}{1+\theta} f(s) ds + \mu \int_{s_0}^{s_1} \left[sy(s) - \frac{y(s)^{1+\epsilon}}{1+\epsilon} - u(s) \right] f(s) ds \\ + \int_{s_0}^{s_1} \lambda(s) \left[u(s_0) + \int_{s_0}^{s} y(w)^{1+\epsilon} w^{-1} dw - u(s) \right] ds\end{aligned} \tag{3.10}$$

This function is to maximized by choice of functions $\{(y(s), u(s)\}$. The solution is by pointwise maximization, with first-order conditions

$$\frac{\partial \mathscr{L}}{\partial u} = (u(s)^{\theta} - \mu) f(s) - \lambda(s) = 0 \tag{3.11}$$

$$\frac{\partial \mathscr{L}}{\partial y} = \mu(s - y(s)^{\epsilon}) f(s) + (1+\epsilon) y(s)^{\epsilon} s^{-1} \Lambda(s) = 0 \tag{3.12}$$

where

$$\Lambda(s) = \int_{s}^{s_1} \lambda(w) dw$$