ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 83/84 – 2006

Tagging and Redistributive Taxation

Robin Boadway*, **Queen's University, Canada**
Pierre Pestieau**, **Université de Liège, Belgique**

ABSTRACT. – We study optimal redistributive taxes when the population can be disaggregated into tagged groups. Under reasonable circumstances, the tax system will be more redistributive in the tagged group with the higher proportion of high-ability persons. We extend the analysis to the case where the tag reflects differences in resources required to achieve a given level of utility. The compensation given for needs depends on whether the income tax structure is differentiated by needs groups.

Ciblage et taxation redistributive

RÉSUMÉ. – Nous étudions le problème de la taxation optimale dans une société qui peut être segmentée en groupes ciblés. Dans des situations plausibles, le système de taxation sera plus redistributif dans le groupe comprenant une proportion relativement plus élevée d'agents à productivité élevée. Ensuite, nous étudions le problème de la taxation optimale lorsque la segmentation concerne des différences de besoins, un groupe devant consommer davantage que les autres pour atteindre un niveau de satisfaction donné. Nous montrons que la compensation accordée pour ces différences de besoins varie selon que la taxation reflète cette segmentation par les besoins.

This paper was presented at the Third Journées d'Économie Publique Louis-André Gérard-Varet, Marseille, June 10-11, 2004. We are grateful to the discussant, Philippe de Donder, to John Weymark and to two referees for helpful comments and suggestions. This research was begun with our colleague and friend, Maurice Marchand, who passed away suddenly in July of 2003.

* R. Boadway: Department of Economics, Queen's University, Kingston, Ontario K7L 3N6, Canada. E-mail: boadwayr@econ.queensu.ca

** P. Pestieau: CREPP – Université de Liège, Boulevard du Rectorat, 7, Bâtiment 31, boite 39, 4000 Liege (Sart-Tilman), Belgique. E-mail: P.Pestieau@ulg.ac.be

1 Introduction

The standard optimal redistributive taxation model assumes that households are distributed over some private characteristic, such as ability, whose distribution is common knowledge. Redistribution policy is constrained by incentive constraints that must be satisfied if households are to reveal their true ability-types (Mirrlees, 1971; Stiglitz, 1982). There are many circumstances in which the population may be disaggregated into groups using some observable characteristic, or 'tag' to use the terminology introduced by Akerlof (1978), which is correlated with ability. For example, households may be identified with the region of their residence, with some demographic characteristic, such as gender or age, or with some physical attribute, such as a health condition or a special need. As long as the distribution of abilities is known to differ within households of different characteristics or tags, this additional information should be useful in enhancing redistribution policy. The government is then able to redistribute both among different ability-types within each tagged group and between groups. While incentive constraints apply within each group, they do not apply between groups since the tag is observable[1].

In this paper, we study various aspects of the optimal redistributive tax structure when the population can be disaggregated into tagged groups. To make the analysis manageable and to obtain qualitative results, we focus on the case where there are two ability types (high- and low-ability) and two tagged groups, although the results can be readily generalized to the multi-type case[2]. We begin with the case in which the tag has no further normative significance: it serves simply to separate the population into identifiable groups, each of which has a distribution of ability-types. We find that under reasonable circumstances, the tax system will be more redistributive in the tagged group that has a higher proportion of high-ability persons. Moreover, inter-group redistribution will always go from the group with a higher proportion of high-ability types to that with a lower proportion. Of course, such a system may be resisted because, if the tagging characteristic has no direct utility consequences, a differentiated tax system violates the principle of horizontal equity.

We then extend the analysis to the case where the tag does have welfare significance. Suppose households vary by needs, where differences in needs refer to differences in the amount of resources required to achieve a given level of utility. This might, for example, be due to a medical condition or a disability. Modeling needs in this way avoids problems of non-comparability of utilities that would arise if needs were reflected in different utility functions. Following Rowe and Woolley (1999)

1. Much of the literature on tagging, also referred to as categorical transfers, simplifies the within group redistribution problem by assuming that incomes are fixed. Examples include Parsons (1996), Keen (1992) and Blackorby and Donaldson (1994). An extension to the case in which labor supply is variable may be found in Salanié (2002).
2. Immonen *et al.* (1998) have studied the pattern of optimal marginal income tax rates in a simulated model with continuous abilities and two tagged groups. Their analysis relies on simulations to obtain suggestive results. Comparative static results on optimal non-linear income tax structures are notoriously scarce. One example can be found in Weymark (1987). He considers the effect on the tax structure of changes in social welfare weights and individual preferences rather than changes in the distribution of abilities, which is our concern. See also the recent paper by Brett and Weymark (2004). Viard (2001a,b) studies tagging in a linear income taxation context in which a common marginal tax rate applies economy-wide and the lump-sum component differs across tagged groups. There are no needs differences between tagged groups.

and Boadway and Pestieau (2003b), we consider the case in which needs reflect differences in consumption requirements[3]. Households now vary by two characteristics — ability and needs — which are assumed to be imperfectly correlated. While ability is private information to the households, needs are observable and can be used as a tag. Inter-group transfers now take on an additional responsibility — that of compensating for differences in needs. In a full-information setting, needs would be fully compensated by lump-sum transfers, and lump-sum redistribution would also apply between ability-types. The same principle extends to the imperfect-information case in a natural way. Again, horizontal equity will typically be violated, and that may be a concern. We consider the level of compensation for needs when the income tax schedule cannot be differentiated by needs group, whether for horizontal equity or other reasons. This is the case considered by Rowe and Woolley (1999). The level of compensation given for needs in this case exceeds the level of needs if a higher proportion of low-ability households are needy, and vice versa.

Strangely enough, the issue at hand, namely optimal taxation with tagging and transfers between tagged groups, has received little attention in the literature. There is the exception of Immonen *et al.* (1998) already mentioned. They explore the design of an optimal income tax when some categorical information is available, and show that the use of such information may entail important gains. Unlike our paper, they mainly resort to simulations to obtain some results. Another paper by Hamilton and Pestieau (2005) explores the incidence of different proportions of skilled-unskilled workers on the income tax schedule and on post-tax utility. From there, they show the effect of interregional mobility of either skilled or non-skilled on social welfare.

Income taxation with tagging raises questions of ethics and of political feasibility. Starting with the first, consider a country with two observable groups, the Greens and the Blues, where it is known that the proportion of skilled workers is much higher among the Greens than among the Blues. Even if we show that differentiating tax policy according to people's color accompanied by transfers from the Greens to the Blues dominates a unique tax system from a welfarist point of view, it may be difficult to implement. As mentioned, segmenting tax policy generates horizontal inequities and violates principles of neutrality, and these may be regarded as undesirable from an ethical point of view[4]. As to the political feasibility, take a country with two regions having a different mix of skilled and unskilled. Again, designing a separate tax schedule for each of the two regions with transfers from the region with the better endowment in human capital to the other one is to be preferred on welfarism grounds. Yet, the mere existence of such transfers underlines the dependence of one region on the other and can lead to separatist, or at least egoistic, reactions. By keeping a single tax schedule these aspects do not appear in the spotlight[5].

3. The alternative case in which needs reflect different leisure requirements is considered in a different context in Boadway and Pestieau (2003a).
4. The use of tagging to separate persons might also be resisted by those being tagged because of the stigma that might be involved. Evidence of less than complete take-up rates in welfare schemes, such as that reported in Hernanz et al. (2004), supports this. For an analysis of the effects of stigmatization on the optimal use of tagged transfers, see Jacquet and van der Linden (2006).
5. The conflict between horizontal equity and social welfare maximization has been recognized in the fiscal federalism literature. The consequences of this conflict for the design of intergovernmental fiscal relations is discussed in Boadway (2004). Of course, it is also the case that if the tag is interpreted as region of residence, the possibility of migration would reduce the benefits of differential tax structures by region as well.

We proceed in the following section by setting up our basic model in the absence of tagging. We then consider the case of tagging where the tag reflects no difference in household utilities. Then we extend the analysis to the case where the tag applies to persons with different needs. A final section concludes.

2 The Basic Model without Tagging

Obtaining analytical results in general optimal income tax models is notoriously difficult, which accounts for the important role that simulations have played. Our strategy is to specify a model that does allow for qualitative results. Obviously this detracts from the generality of the results, but it does serve to highlight the intuition and some of the main influences at work. Our model specializes the standard Mirrlees-Stiglitz optimal income tax model in two main ways. First, we adopt a quasi-linear-in-consumption additive specification for household utility. This utility function, also used by Diamond (1998), essentially eliminates income effects from labor supply. Second, the social welfare function we use exhibits constant *absolute* aversion to inequality, as opposed to the more standard constant relative aversion to inequality form. This implies that social indifference curves, instead of being homothetic, are parallel in a 45° direction. Other properties of social preferences are preserved, such as the Pareto property, symmetry in household utilities and, to ensure non-negative aversion to inequality, concavity[6]. Otherwise, the standard assumptions of optimal income taxation apply.

Following Stiglitz (1982), we restrict the number of household ability-types to two, although as we shall indicate how the results can readily be extended to the multi-type case as in Guesnerie and Seade (1982). Households may be one of two ability levels, high-ability w_2 and low-ability w_1, with $w_2 > w_1$, where w_i corresponds with the wage rate of a type-*i* household. The proportion of low-ability households in the population is given by π, so $1-\pi$ are high-ability. For simplicity, we normalize total population to unity.

Initially, we assume that all households have the same preferences, given by the quasi-linear form $c-h(\ell)$, with $h'(\ell),\ h''(\ell)>0$, where c is composite consumption and ℓ is labor. We can regard this as an ordinal utility function, leaving measurability as a matter for social judgment by the government/planner. Since income is given by $y = w_i\ell$, we can rewrite our utility metric in terms of before- and after-tax income for a type-*i* household as:

$$x_i \equiv c - h\left(y/w_i\right)$$

6. The social choice foundations of this social welfare function are discussed in Bossert and Weymark (2002), who refer to it as the symmetric Kolm-Pollak social welfare ordering, after Kolm (1969) and Pollak (1971). It requires that household utility functions satisfy translation-scale measurability and be fully comparable between households.

In what follows, we refer to x_i as the *real income* of a type-*i* household.

Social utility for a type-*i* household, denoted by $u(x_i)$, is given by $u(x_i) = -e^{-\rho x_i}$, with $0 \le \rho \le \infty$. Social welfare is then the sum of social utilities:

$$w(x_1, x_2) = \pi u(x_1) + (1-\pi) u(x_2) = -\pi e^{-\rho x_1} - (1-\pi) e^{-\rho x_2} \tag{1}$$

Note that $\rho = -u''(x)/u'(x)$. This can be interpreted as the *coefficient of absolute inequality aversion*. Given the definition of social welfare in (1), the slope of a social indifference curve in real income space is:

$$\left.\frac{dx_1}{dx_2}\right|_w = -\frac{1-\pi}{\pi}\frac{e^{-\rho x_2}}{e^{-\rho x_1}} = -\frac{1-\pi}{\pi} e^{-\rho(x_2 - x_1)}$$

This implies, as noted, that social indifference curves are parallel in a 45° direction. For $\rho = 0$, we have the utilitarian case with linear social indifference curves: $w(\cdot) = \pi x_1 + (1-\pi) x_2$. For $\rho = \infty$, we have maximin.

2.1 The Income Possibilities Frontier

A useful pedagogical device in this simple economy is the Pareto frontier in household real income space, referred to as the second-best *Income Possibilities Frontier* or IPF. A point on the IPF where the incentive constraint on high-ability households is binding (that is, one that would be chosen if aversion to inequality is non-negative) is the solution to the following Pareto-optimizing problem:

$$(\mathcal{P}) \qquad \max_{\{c_i, y_i\}} \; x_1 = c_1 - h\left(\frac{y_1}{w_1}\right)$$

subject to

$$(\mu) \qquad c_2 - h(y_2 / w_2) - m \ge 0$$

$$(\gamma) \qquad c_2 - h(y_2/w_2) - c_1 + h(y_1/w_2) \ge 0$$

$$(\lambda) \qquad \pi(y_1 - c_1) + (1-\pi)(y_2 - c_2) + R \ge 0$$

In this problem, *m* is the predetermined real income level of type-2 households; the second constraint is the incentive constraint applying to the high-ability types; and the third constraint is the resource constraint, assuming for simplicity that there is an exogenously given amount of revenue *R* available. The equation labels refer to the respective Lagrangian multipliers for the constraints. Thus, the Lagrangian expression is:

$$\mathcal{L} = c_1 - h(y_1/w_1) + \mu\left[c_2 - h(y_2/w_2) - m\right] + \gamma\left[c_2 - h(y_2/w_2) - c_1 + h(y_1/w_2)\right]$$
$$+\lambda\left[\pi(y_1 - c_1) + (1-\pi)(y_2 - c_2) + R\right]$$

The first-order conditions reduce to the following:

(2) $$\mu = \lambda - 1, \quad \gamma = 1 - \lambda\pi$$

(3) $$h'(y_2/w_2) = w_2$$

(4) $$(1-\lambda\pi)h'(y_1/w_2)/w_2 + \lambda\pi - h'(y_1/w_1)/w_1 = 0$$

(5) $$\pi\left(y_1 - h(y_1/w_2) - m\right) + (1-\pi)\left(y_2 - h(y_2/w_2) - m\right) + R = 0$$

Equation (3) is the familiar zero-marginal-tax-rate-at-the-top condition. An advantage of our assumption about preferences is that (3) pins down y_2, so that we can take y_2 as given in what follows. Then, (4) and (5) determine y_1 and λ in terms of m, R and π [7].

More generally, the solution to the above problem yields the value function $x_1(m, R, \pi)$. Given that $m = x_2$, this is just the expression for the IPF, given R and π. By the envelope theorem, we have:

(6) $$\frac{\partial x_1}{\partial m} = -\mu = 1 - \lambda < 0, \quad \frac{\partial x_1}{\partial R} = \lambda > 0,$$
$$\frac{\partial x_1}{\partial \pi} = \lambda\left[(y_1 - c_1) - (y_2 - c_2)\right] = \lambda\left[t_1 - t_2\right] < 0$$

where t_i is the tax paid by a type-i person. Thus, $-\mu = 1 - \lambda$ is the slope of the IPF.

The shape of the IPF can be determined by investigating how λ changes with m, the real income given for type-2 households. Note first that if the incentive constraint on type-2 households is not binding, $\gamma = 0$. From the first-order conditions on (c_1) and (c_2), we have $-\mu = -(1-\pi)/\pi$, so the IPF is linear. Suppose next that

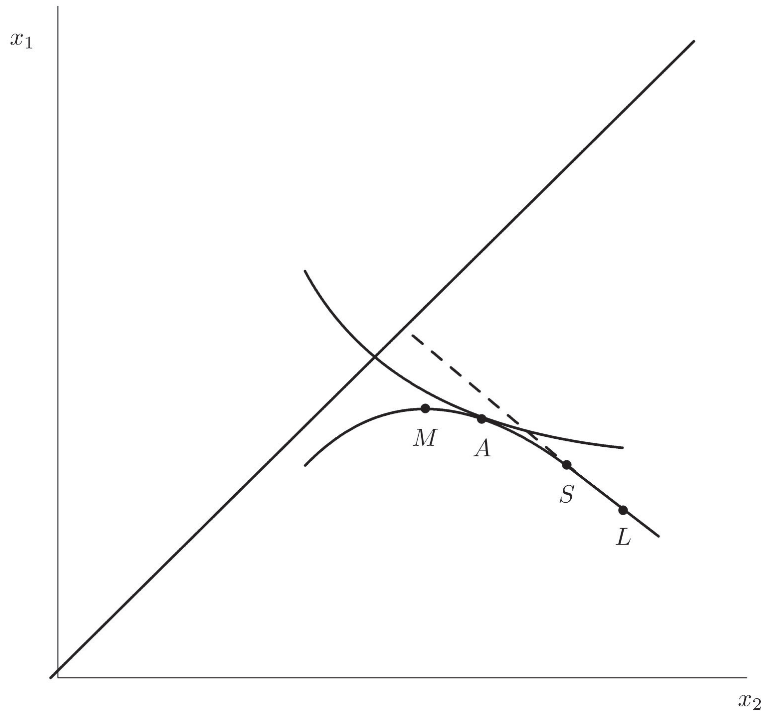
7. We assume in what follows that $y_1 > 0$ in the optimum so that we have an interior solution. A sufficient condition for this is that $h'(0)$ be arbitrarily small.

the incentive constraint is binding. Differentiating (4) and (5), we obtain the following comparative static result:

$$\frac{\partial\lambda}{\partial m} = -\frac{(1-\lambda\pi)h''(y_1/w_2)/w_2^2 - h''(y_1/w_1)/w_1^2}{\left[\pi(1-h'(y_1/w_2))/w_2\right]^2} = \frac{d\mu}{dm}$$

In what follows we assume that $\partial\mu/\partial m > 0$, so that the IPF is strictly concave. This will be the case if $(1-\lambda\pi)h''(y_1/w_2)/w_2^2 - h''(y_1/w_1)/w_1^2 / < 0$, which is just the second-order condition for a maximin optimum[8]. More generally, the condition will be satisfied if $h''(y_1/w_2) \le h''(y_1/w_1)$, or $h'''(y/w) \ge 0$. A quadratic labor cost function will be sufficient for this inequality to hold.

FIGURE 1
Income Possibilities Frontier (IPF)



8. The maximin problem is $\max\ c_1 - h(y_1/w_1)$ subject to conditions (γ) and (λ). Substitution from the contraints makes this equivalent to max $\pi y_1 + (1-\pi)\left[y_2 + h(y_1/w_2)(y_2/w_2)\right] + R - h(y_1/w_1)$ with respect to y_1 (since y_2 is determined by (3)). The second-order condition for this problem is as stated.

Figure 1 illustrates. The curve extending through L, S, A and M illustrates a typical IPF. The point L is the laissez-faire allocation. The IPF is linear until the point S, where the incentive constraint is binding. The point M is the maximin point, considered further below. With no aversion to real income inequality $(\rho = 0)$, points along the segment LS are all socially optimal. For positive but finite aversion to inequality the optimal point will lie between S and M, such as at the point A. The first-best IPF is shown as a dashed line. Since the incentive constraint is not binding in the first best (because of full information), the slope is $-\mu = -(1-\pi)/\pi$ throughout. Recall that the slope of social indifference curves will also be $-(1-\pi)/\pi$ in the utilitarian case where $\rho = 0$. In this case, the laissez-faire is socially optimal, and there is no case for policy intervention. In what follows, we therefore assume that $\rho > 0$.

The IPF will shift with a change in either π or R, both of which are of interest to us. Consider each in turn.

2.2 Effect of Changes in π on the IPF

Changes in π cause the entire IPF to shift in (x_2, x_1)– space. An indication of the effects of changing π on the IPF can be obtained in principle by solving for $\partial\lambda/\partial\pi$ from (4) and (5) above. However, the resulting expression is of ambiguous sign so offers little insight. It is more useful to characterize shifts in the IPF with reference to particular optimal allocations.

Consider the planning problem for a given specification for social utility, $u(x_1)$, that is, for a given choice of the coefficient of absolute inequality aversion, ρ. Using the definition of real income x_i, the second-best planning problem can be written:

$$(\mathcal{A}) \qquad \max \quad \pi u(x_1) + (1-\pi) u(x_2)$$

subject to

$$(\gamma) \qquad x_2 - x_1 + h(y_1/w_2) - h(y_1/w_1) \geq 0$$

$$(\lambda) \qquad \pi\big(y_1 - x_1 - h(y_1/w_1)\big) + (1-\pi)\big(y_2 - x_2 - h(y_2/w_2)\big) + R \geq 0$$

where the constraints (γ) and (λ) are as before the incentive constraint and the government revenue constraint. The first-order conditions reduce, after substituting out the Lagrangian multipliers, to the following equation:

$$(7) \qquad \begin{aligned} &(1-\pi)\big[u'(x_1) - u'(x_2)\big]\big[h'(y_1/w_2)/w_2 - h'(y_1/w_1)/w_1\big] \\ &\qquad + \big[\pi u'(x_1) + (1-\pi)u'(x_2)\big]\big[1 - h'(y_1/w_1)/w_1\big] = 0 \end{aligned}$$

This equation, along with constraints (γ) and (λ), constitute three equations in x_1, x_2 and y_1 (where y_2 is again determined by the no-distortion-at-the-top condition (3)).

Differentiating these three conditions, we can derive the following results:

$$D\frac{dx_1}{d\pi} = (t_1 - t_2)\left[C_{y_1} - C_{x_2}(\theta_{21} - \theta_{11})\right] - C_\pi\left[(1-\theta_{11})\pi + (\theta_{21} - \theta_{11})(1-\pi)\right] > 0 \tag{8}$$

$$D\frac{dx_2}{d\pi} = (t_1 - t_2)\left[C_{y_1} + C_{x_1}(\theta_{21} - \theta_{11})\right] - C_\pi\left[(1-\theta_{11})\pi - (\theta_{21} - \theta_{11})\pi\right] \gtrless 0 \tag{9}$$

where

$$\theta_{11} \equiv h'(y_1/w_1), \qquad \theta_{21} \equiv h'(y_1/w_2)/w_2$$

$$C_{x_1} = (1-\pi)\left[(\theta_{21} - \theta_{11}) + \pi(1-\theta_{21})\right]u''(x_1) \gtrless 0$$

$$C_{x_2} = (1-\pi)(1-\theta_{21})u''(x_2) < 0$$

$$C_{y_1} = (1-\pi)\left(u'(x_1) - u'(x_2)\right)\left[h''(y_1/w_2)/w_2^2 - h''(y_1/w_1)/w_1^2\right]$$

$$-\left[\pi u'(x_1) + (1-\pi)u'(x_2)\right]h''(y_1/w_1)/w_1^2 < 0$$

$$C_\pi = \left(u'(x_1) - u'(x_2)\right)(1-\theta_{21}) > 0$$

and $D < 0$ is the determinant of the matrix of derivatives of the endogenous variables. Its sign and that of C_{y_1} are determined by the second-order conditions, which we assume are satisfied.

Combining (8) and (9) and using $D < 0$ and $\theta_{21} - \theta_{11} < 0$, we obtain:

$$\frac{dx_1}{d\pi} < \frac{dx_2}{d\pi} \text{ if } (t_2 - t_1)\left(C_{x_2} + C_{x_1}\right) - C_\pi < 0$$

that is, if

$$(10)\qquad \begin{gathered}(t_2-t_1)\left[u''(x_1)(\theta_{21}-\theta_{11})+(\pi u''(x_1)+(1-\pi)u''(x_2))(1-\theta_{21})\right]\\ -(u'(x_1)-u'(x_2))(1-\theta_{21})<0\end{gathered}$$

It is straightforward to show that this condition is satisfied when the social utility function exhibits constant absolute inequality aversion: $u(x_i)=-e^{-\rho x_i}$. To see this, substitute $\rho=-u''(x_i)/u'(x_i)$ into (7) to obtain:

$$u''(x_1)(\theta_{21}-\theta_{11})+\left[\pi u''(x_1)+(1-\pi)u''(x_2)\right]\left[1-\theta_{21}\right]=0$$

Using this, (10) becomes:

$$-(u'(x_1)-u'(x_2))(1-\theta_{21})=-C_\pi<0$$

Therefore,

$$\frac{dx_1}{d\pi}<\frac{dx_2}{d\pi}$$

for any value of $\rho>0$. As well, $dx_1/d\pi<0$, while $dx_2/d\pi<0$.

Figure 2 illustrates the effect on the IPF of an increase in the proportion of low-ability households, π. The laissez faire point L remains unchanged, while the point S at which the incentive constraint is just binding moves inward along a 45° line. The maximin point M, where $\rho=\infty$, and the point A, where $0<\rho<\infty$, move in a direction that is steeper than 45°[9].

---

9. If social utility exhibits constant relative aversion to income inequality $\left(u(x_i)=x_i^{1-\rho}/(1-\rho)\right)$, Figure 2 may not apply. In this case, using $\rho=-u''(x_i)x_i/u'(x_i)$ and (7), the lefthand side of (10) becomes:

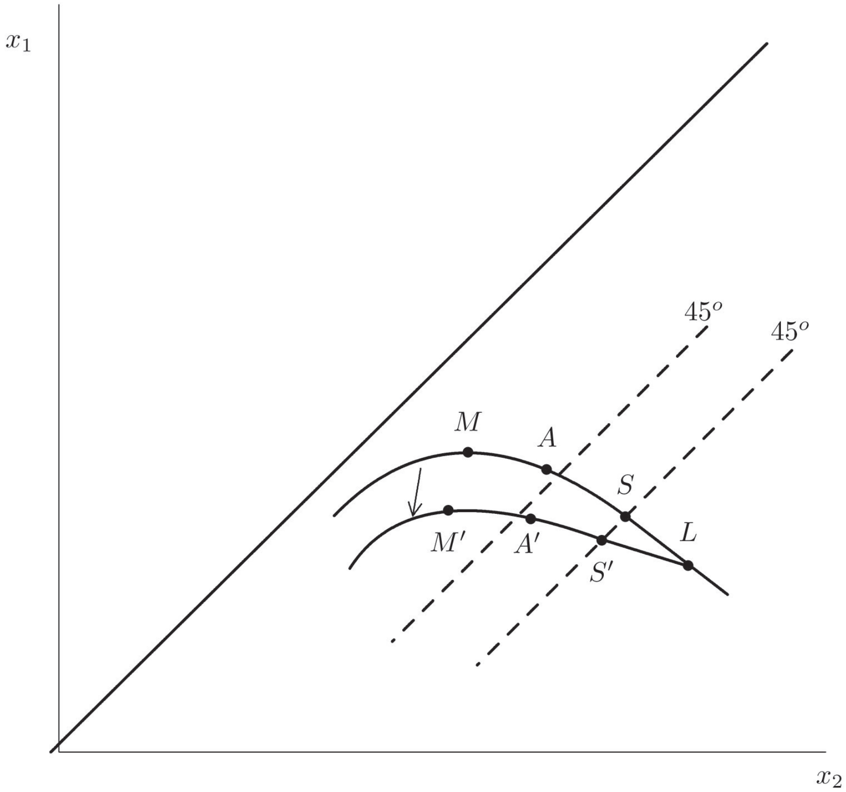
$$(t_2-t_1)(1-\pi)u''(x_2)(1-x_2/x_1)-(u'(x_1)-u'(x_2))$$

Since the terms being subtracted are both positive, (10) is not necessarily satisfied so we cannot say unambiguously how x_1 changes relative to x_2 as π changes. However, as $\rho\to\infty$, condition (10) will be satisfied. Substitute $u''(x_2)=-u'(x_2)\rho/x_2$ into the above to get:

$$-(t_2-t_1)(1-\pi)\rho(1-x_2/x_1)/x_2-\left((x_2/x_1)^\rho-1\right)$$

As ρ increases, $(x_2/x_1)^\rho$ goes to ∞ more rapidly than ρ goes to ∞ by l'Hôpital's Rule. Therefore, expression (10) becomes negative for ρ large enough.

FIGURE 2
Effect of* π *Increase on IPF



2.3 Effect of Changes in *R* on the IPF

An increase in government revenue availability *R* causes the entire IPF to shift outwards in a 45° direction. To see this, consider a point along the IPF where the self-selection constraint is binding. In (y,c) – space, it can be characterized as a standard Stiglitz-type separating equilibrium with indifference curves for ability-types 1 and 2, where type 2's indifference curve intersects type 1's at the latter's bundle (y_1,c_1). Since utility is quasi-linear in *c*, indifference curves are parallel vertically.

Consider now an increase in revenue ΔR. If both c_1 and c_2 increase by ΔR, holding y_1 and y_2 constant, all equilibrium conditions are still satisfied, including the budget constraint, the self-selection constraint, and the tax structure. Therefore, a change in *R* causes the entire IPF to shift in a 45° degree direction.

We are now in a position to consider the effects of tagging. We begin with the case in which tagging has no welfare consequences. It simply separates households into two groups with different ability distributions. Subsequently, we take up the case where tagging corresponds with a difference in needs between tagged and untagged persons.

3 Tagging with No Needs Differences

Suppose the population can be divided into two groups with different proportions of ability-type 1's, $\overline{\pi} > \underline{\pi}$. For simplicity, assume the groups are of equal size, so $\overline{\pi} + \underline{\pi} = 2\pi$. In this section, the only thing that distinguishes the two groups is the skill distribution of the population. For example, the two groups could be treated as separate regions within the same country. In this case, our discussion above of the effect of changes in π on the IPF can be applied directly. These two groups will have IPF's that bracket that of the two groups taken together — the pooling IPF — as shown in Figure 3. Note that the points S, $\underline{S}$ and $\overline{S}$ where the self-selection constraints are just binding all lie along a 45° line, while the maximin points M, $\underline{M}$ and $\overline{M}$ and the social utilitarian points A, $\underline{A}$ and $\overline{A}$ each line up along a locus that is steeper than 45°. For any value of $\rho > 0$, type 1's are unambiguously worse off in the tagged group that has a higher proportion of low-ability types $\left(\overline{\pi}\right)$, while type 2's could be better off or worse off. In that sense, we can say that the tax structure is more progressive the lower the proportion of low-ability types there are in the population. Moreover, tagging generally makes some of both ability types worse off than pooling. This characterization follows directly from Figure 3.

The IPFs for the tagged groups in Figure 3 involve only intra-group redistribution. Since the government is able to observe which persons are in each group, lump-sum inter-group transfers will be possible. We can now show that such transfers will be desirable, and will go from the $\underline{\pi}$ to the $\overline{\pi}$ group.

To show this, we establish first that any point along the pooling IPF can be replicated by a set of transfers $\left(\overline{T}, \underline{T}\right)$ between the two tagged groups combined with the same common optimal income tax structure in each group. To see this, note that the government budget in the pooling equilibrium is:

$$2\pi t_1 + 2\left(1-\pi\right)t_2 = 0$$

where t_1 and t_2 are the taxes levied on the two ability types. To finance the same pooling tax schedule in the two tagged groups, the following budgets must be satisfied:

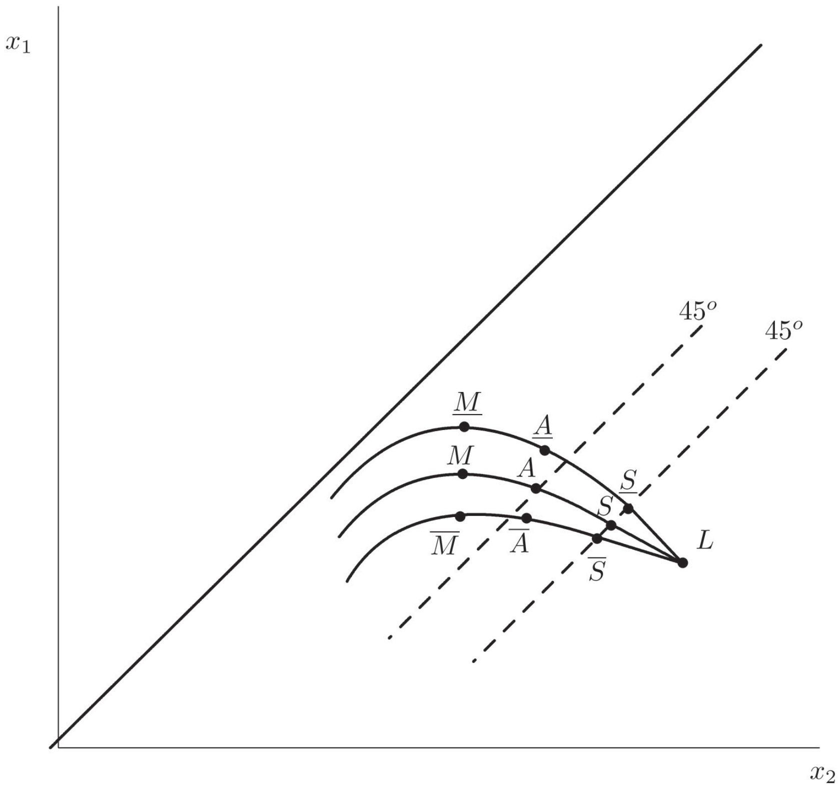
$$\overline{\pi} t_1 + \left(1-\overline{\pi}\right)t_2 = \overline{T} \text{ and } \underline{\pi} t_1 + \left(1-\underline{\pi}\right)t_2 = \underline{T}$$

Adding these together, we obtain:

$$\overline{T} + \underline{T} = \left(\overline{\pi} + \underline{\pi}\right)t_1 + \left(1-\overline{\pi}+1-\underline{\pi}\right)t_2 = 2\pi t_1 + 2\left(1-\pi\right)t_2 = 0$$

by the pooled budget constraint. Recall from above that an increase in revenues R given to a group causes its IPF to shift outward in a 45° direction, and vice versa. Therefore, any given point on the pooled IPF can be disaggregated into intersecting

FIGURE 3
IPFs for Tagged and Pooled Groups



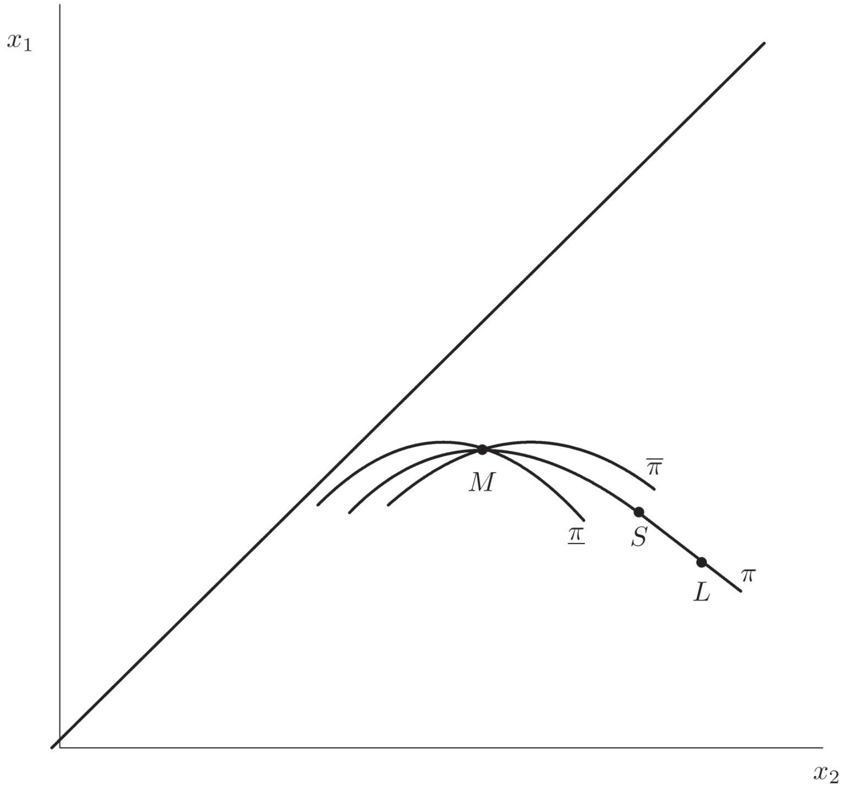
IPFs for the two tagged groups by a transfer from group $\underline{\pi}$ to group $\overline{\pi}$. We expect that at the point of intersection, the IPF for the $\underline{\pi}$ group (whose IPF is shifted down) will be steeper than that for the $\overline{\pi}$ group (whose IPF is shifted up).

These effects of inter-group transfers can be used to characterize optimal tagging outcomes for given specifications of the government's objectives. It is useful to discuss separately the maximin case and the case of finite aversion to inequality.

3.1 The Maximin Optimum with Tagging

The maximin outcome on the pooled IPF is the point *M* in Figure 4. This point can be achieved by a set of lump-sum transfers $\overline{T}$ and $\underline{T}$ between the tagged groups, combined with applying the pooled maximin solution tax structure in each tagged group. The figure also indicates the IPFs for the two groups $\overline{\pi}$ and $\underline{\pi}$ given the inter-group transfers. Note that at the point *M*, the IPF for $\underline{\pi}$ is steeper than that for $\overline{\pi}$. This follows from the facts that i) the transfers $\overline{T}$ and $\underline{T}$ will shift the two IPFs in a 45° direction and ii) the points $\underline{M}$ and $\overline{M}$ in Figure 3 are lined up along a locus that is steeper than 45°. Since the intersection point is

FIGURE 4
The Pooled Maximin Outcome



at a peak for the pooled case, this implies that the peaks for the two groups are on either side of that for the pooled case, and unambiguously above it vertically. The full maximin will generally require further redistribution from one group to another since the low-ability persons will generally have different real income levels in Figure 3. Nonetheless, both persons will end up being better off than in the pooled case.

The following results are apparent for the maximin case:

1. The worst-off person in both groups can be made better off by moving from the pooled maximin case to two separate maximin outcomes with different tax schedules and inter-group lump-sum transfers.

2. In the full maximin optimum with two separate tax schedules, the type-2's are worse off in tagged group $\underline{\pi}$, that is, the one with a lower proportion of low-ability types. In other words, the tax is more progressive in the tagged group with the highest proportion of high-ability types.

For future reference, note that the optimal tagging outcome with two separate tax schedules violates the principle of horizontal equity. In this maximin case, high-ability types are treated differently by the tax system depending solely on which group they find themselves in. If horizontal equity is an objective that must be satisfied, only the pooled outcome would be possible.

3.2 Tagging with Finite Absolute Aversion to Inequality

In this sub-section, we focus on outcomes for an aversion to inequality parameter ρ such that $0 < \rho < \infty$. This corresponds with point A in Figure 2. As the figure indicates, as π increases, point A moves southwest along a locus that is steeper than 45° (and either positive or negative). Point A' is one point along that locus. Figure 3 uses this finding to depict how real income combinations will deviate from the pooled optimum when households are separated into tagged groups with different values of π, assuming there are no inter-group transfers. As the figure indicates, social welfare will be higher in the group with the lowest proportion of low-ability types, $\underline{\pi}$. Moreover, tagging in the absence of inter-group transfers will reduce social welfare in the $\overline{\pi}$ group. This suggests that there will be some scope for inter-group transfers. It turns out that tagging combined with lump-sum inter-group transfers can make social welfare in both groups higher that it is in the pooling optimum[10]. As well, progressivity is higher in the tagged group with the lowest proportion of low-ability types.

To see this, we first show that, starting in the tagged outcomes $\left(\underline{A}, \overline{A}\right)$ depicted in Figure 3 with no inter-group transfers, lump-sum redistribution should go from the $\underline{\pi}$-group to the $\overline{\pi}$-group. Denote by $\overline{x}_i$ and $\underline{x}_i$ the levels of real income obtained by type-i households in the two tagged groups. In the absence of inter-group transfers, social welfare levels satisfy $w\left(\overline{x}_1, \overline{x}_2\right) < w\left(\underline{x}_1, \underline{x}_2\right)$ since social welfare is decreasing in π by the envelope theorem applied to problem (A). Using (1), this implies:

$$-\overline{\pi}e^{-\rho\overline{x}_1} - \left(1-\overline{\pi}\right)e^{-\rho\overline{x}_2} < -\underline{\pi}e^{-\rho\underline{x}_1} - \left(1-\underline{\pi}\right)e^{-\rho\underline{x}_2} \tag{11}$$

We know from above that a change in R causes c_1 and c_2 to change by the same amount, with y_1 and y_2 unchanged. Therefore, x_1 and x_2 also change by dR, so we have:

$$\frac{dw\left(\overline{x}_1, \overline{x}_2\right)}{dR} = \rho\left[\overline{\pi}e^{-\rho\overline{x}_1} + \left(1-\overline{\pi}\right)e^{-\rho\overline{x}_2}\right], \quad \frac{dw\left(\underline{x}_1, \underline{x}_2\right)}{dR} = \rho\left[\underline{\pi}e^{-\rho\underline{x}_1} + \left(1-\underline{\pi}\right)e^{-\rho\underline{x}_2}\right] \tag{12}$$

Equation (11) implies:

$$\rho\left[\overline{\pi}e^{-\rho\overline{x}_1} + \left(1-\overline{\pi}\right)e^{-\rho\overline{x}_2}\right] > \rho\left[\underline{\pi}e^{-\rho\underline{x}_1} + \left(1-\underline{\pi}\right)e^{-\rho\underline{x}_2}\right]$$

10. Of course, aggregate social welfare will necessarily rise when tagging combined with inter-group transfers is used, regardless of the form of individual and social utility functions, as a referee has pointed out. The no-tagging outcome is always a feasible choice when tagging can be used. If the optimal outcome involves tagging with different tax structures in the tagged groups, social welfare will be higher by revealed preference. More formally, in the no tagging case, social welfare is a convex function of the distribution of ability types, say, $V(F)$. When tagging allows the distribution to be divided into two sub-distributions F_1, F_2 in the proportions λ_1, λ_2 convexity of social welfare $V(\cdot)$ implies $V(F) \le \lambda_1 V(F_1) + \lambda_2 V(F_2)$, where the equality only applies if the optimal tax schedule is the same for F_1 and F_2 (e.g., if the two distributions are of the same form).

so

$$\frac{dw(\overline{x}_1,\overline{x}_2)}{dR} > \frac{dw(\underline{x}_1,\underline{x}_2)}{dR}$$

Therefore, when the population is separated into two groups, the planner will want to redistribute lump-sum from the $\underline{\pi}$-group to the $\overline{\pi}$-group until: $dw(\overline{x}_1,\overline{x}_2)/dR = dw(\underline{x}_1,\underline{x}_2)/dR$ or,

$$\overline{\pi}e^{-\rho\overline{x}_1} + (1-\overline{\pi})e^{-\rho\overline{x}_2} = \underline{\pi}e^{-\rho\underline{x}_1} + (1-\underline{\pi})e^{-\rho\underline{x}_2}$$

Consider now the effect of redistributing between the two groups. We know that we can redistribute between groups to achieve the common pooled solution A. At the common pooled solution, the IPFs for the two groups will intersect since the redistribution causes the curves to move in a 45° direction. Point $\underline{A}$ on $\underline{\pi}$'s new IPF will be to the left of A and on a higher social indifference curve, while $\overline{A}$ on $\overline{\pi}$'s new IPF will be to the right of A and also on a higher social indifference curve, as shown in Figure 5. Within each tagged group, social welfare can then be increased by moving from the pooled tax structure at A to group-specific tax structures at $\underline{A}$ and $\overline{A}$. Then, there could be further inter-group redistribution between $\underline{\pi}$ and $\overline{\pi}$, but both groups will end up being better off than at A. Thus, tagging is unambiguously welfare improving.

These results parallel those obtained in the maximin case. They can be summarized as follows:

1. Social welfare within each tagged group can be improved compared with the pooling outcome by making lump-sum transfers from group $\underline{\pi}$ to group $\overline{\pi}$ and adopting a separate income tax structure within each tagged group.

2. The income tax structure will be more progressive in terms of real income in the group with the lower proportion of low-ability households: $\underline{x}_1/\underline{x}_2 > \overline{x}_1/\overline{x}_2$.

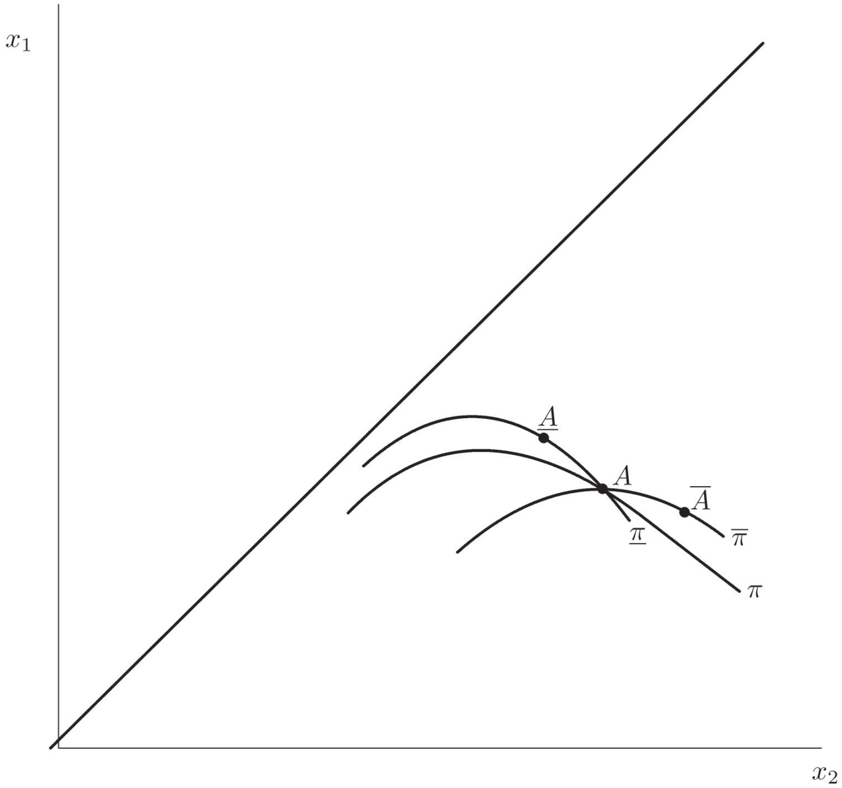
3. Horizontal equity will be violated in the tagging solution in the sense that otherwise identical households will be treated differently depending on which tagged group they belong to.

3.3 Extensions

Intuition suggests that these results can be extended to more complicated settings. The extension to more than two tagged groups is straightforward. The groups can be ordered according to the proportion of the population that are type-1's. Some groups will have higher values of π than the pooled population, and some less. The analog of Figure 5 can be constructed in which the slopes of IPF curves passing through the pooling outcome will be increasing in π. Combining lump-sum inter-group transfers with group-specific tax structures will increase social welfare within each group compared with the pooled outcome. And, the progressivity of group tax structures will be falling in π.

There could also be more than two ability types. Although the analysis of optimal income taxation generalizes in a relatively straightforward way (Guesnerie

FIGURE 5
The Pooled Outcome with Constant Absolute Aversion to Inequality



and Seade, 1982), tagging becomes more complicated. Two conceptual issues arise. First, the comparison of ability distributions across tagged groups becomes ambiguous unless the densities satisfy single-crossing attributes. Second, tagging may result in different groups only having subsets of ability-types. Intuition suggests that if the density of ability distributions across pairs of groups crossed only once, and if all ability-types were in all groups, the analog of the above results should apply. Social welfare within each group should be increased by inter-group transfers and group-specific tax structures, and tax schedules should be more progressive in groups that have more skill-intensive ability distributions.

We have already indicated that the results do not generalize unambiguously to the case of social utility functions that exhibit constant relative inequality aversion. Moreover, they will not extend to more general formulations of household preferences. In these cases, qualitative results will not be possible, though presumably the intuition of the above results will still apply.

4 Tagging with Needs Differences

In this section, we assume that tagging is associated with differences in household needs for resources to achieve a given utility level. These needs differences could reflect a disability or medical condition that can be addressed by a transfer of income. To keep matters simple and tractable, we assume that households can be of two needs types: the needy and the non-needy. All needy households regardless of their ability require an additional amount of consumption goods *n* to obtain comparable utility to the non-needy. This formulation of needs was introduced into the optimal redistribution literature by Rowe and Woolley (1999). Notice that it is analogous to Stone-Geary preferences. Boadway and Pestieau (2003a, b) also consider the effects for redistribution of differences in needs for leisure. In this formulation, unlike with consumption needs, the cost of needs is increasing in the wage rate. Since the government cannot observe wages, this makes optimal policy more complicated. Given our quasi-linear in consumption formulation, preferences for the non-needy are as before $c-h(y/w_i)\equiv \underline{x}_i\ (i=1,2)$, while those for the needy are $c-n-h(y/w_i)\equiv \overline{x}_i\ (i=1,2)$. This quasi-linear formulation constrains our analysis besides simplifying it. It constrains it because preference orderings in (y,c)-space for households of given ability but different needs are identical. Therefore, the single-crossing property does not apply and they cannot be separated. Nonetheless, social utility will differ for the two households: $u(c-y/w_i)>u(c-n-y/w_i)$. At the same time, this formulation of needs simplifies our analysis by avoiding multi-dimensional screening problems.

The proportion of the population who are needy is ϕ, so $1-\phi$ are non-needy. Within the needy group, a proportion $\overline{\pi}$ are low-ability (type 1), while within the non-needy group the proportion is $\underline{\pi}$. In principle, $\overline{\pi} \gtreqless \underline{\pi}$, but special attention will be paid to the case where needs are negatively correlated with ability, so $\overline{\pi}>\underline{\pi}$.

As a benchmark, it is useful to characterize the first-best outcome in which the planner can observe both the ability and the needs of each household. Lump-sum transfers can be made both between needs groups and between ability groups. For any strictly positive value of aversion to inequality ρ, differences in needs *n* will be fully compensated for regardless of ability: $u(\overline{x}_i)=u(\underline{x}_i)$, $(i=1,2)$. Then, lump-sum transfers will be made between ability groups depending on the extent of aversion to inequality. Given our assumption that household preferences are quasi-linear, utilities will be equalized for any positive aversion to inequality $(\rho>0)$. This might be contrasted with the celebrated finding of Mirrless (1974) that if leisure is a (strictly) normal good and household utility functions are strictly concave, utility is declining with ability under utilitarianism. With preferences that are quasi-linear in consumption, leisure is not strictly normal: its income elasticity of demand is zero. Therefore, in the first-best optimum with positive aversion to inequality, equality of real incomes for all households prevails: $\overline{x}_1=\overline{x}_2=\underline{x}_1=\underline{x}_2$.

When abilities are not observable, the first-best cannot be achieved. In fact, the incentive constraint precludes low-ability types from reaching the level of utility of high-ability types. The problem is analogous to the Mirrlees-Stiglitz optimal income tax problem except that differences in needs must be accounted for. We proceed by discussing the tagging case in which needy and non-needy groups can be subject to separate income tax schedules, and then consider the case in which needs must be accounted for in a common income tax schedule, say, because of horizontal equity considerations.

4.1 Tagging with Needs

When the needy and non-needy groups can be separated, the analysis of the previous section can be adopted in a straightforward way. The redistribution system will include both distortionary transfers from the high- to the low-ability persons within each of the needy and non-needy groups, and lump-sum transfers from the non-needy to the needy group. The latter will simultaneously take account of differences in need and differences in the proportion of high-and low-ability types.

Conceptually, inter-group transfers can be separated into the two types. Lump-sum transfers from the non-needy to the needy group will fully compensate for needs differences. Let b denote transfers to compensate for need. If $\overline{b}$ and $\underline{b}$ are the transfers to the needy and non-needy groups respectively, and if they are self-financing, they will satisfy $\overline{b} - \underline{b} = n$ (full compensation for needs) and $\phi\overline{b} + (1-\phi)\underline{b} = 0$ (budget balance). These imply that $\underline{b} = \phi_n$. This needs-compensating transfer applies regardless of the degree of aversion to inequality ρ, as long as it is strictly positive.

Once needs differences have been addressed by lump-sum transfers, the problem becomes exactly analogous to that of the previous section. A lump-sum transfer must be made from the group with the higher proportion of high-ability types to the other group, and separate tax structures applied in the two groups. Both the size of the transfer and the structure of the redistributive income tax will depend on the degree of aversion to inequality ρ. For any value of $\rho > 0$, the tax will be more progressive in the group with the higher proportion of high-ability persons.

The direction and overall size of the inter-group transfer depends on the correlation between needs and abilities. If needs are inversely correlated with abilities, $\overline{\pi} > \underline{\pi}$: the needy group will have the lower proportion of high-ability types. The transfer from the non-needy to the needy to compensate for needs will reinforce the transfer to compensate for ability levels, and the qualitative analysis of the previous section will apply. On the other hand, in the perhaps less-likely event that needs are correlated with ability $(\overline{\pi} < \underline{\pi})$, the needs transfer will counteract the ability-level inter-group transfer, and the overall transfer could in principle go in either way.

In any event, the tagging outcome will violate horizontal equity in the sense that household of identical real incomes will be treated differently under the income tax system. Alternatively, there may be reasons of a political or constitutional reason why different tax schedules cannot be applied to different groups. The following case examines how needs might be accounted for when a common income tax system must apply.

4.2 **Accounting for Needs without Tagging**

The above discussion suggests that needs should be fully compensated for if a different tax schedule can be applied to needy and non-needy groups. Will this still be the case if a common tax schedule must be applied? Rowe and Woolley (1999) had argued that it should be. In this subsection, we show that imperfect needs compensation should generally be used if the planner is restricted to a common tax schedule.

Let $\overline{b}_1$ and $\overline{b}_2$ be needs credits (transfers) given to needy households of ability types 1 and 2, respectively. When $\overline{b}_1 \neq \overline{b}_2$, these can be thought of as income-tested components of the income tax system. We can set $\underline{b} = 0$ for the non-needy with no loss of generality. If the planner is constrained to impose a common income tax system to the two groups, the policy instruments are income tax variables (c_1, y_1, c_2, y_2), which are common to both needy and non-needy households, as well as $\overline{b}_1$ and $\overline{b}_2$. It is useful to formulate the planner's problem in two conceptual stages. In the first stage, needs credits $\overline{b}_1$ and $\overline{b}_2$ are chosen, and in the second stage the optimal income tax is chosen, given the requirement to finance the credits. We can solve the planner's problem in reverse order.

Given $\overline{b}_1$ and $\overline{b}_2$, the planner's optimal income tax problem is the following one:

$$(\mathcal{N}) \qquad \begin{aligned} &\max_{\{c_i, y_i\}} \phi\left\{\overline{\pi} u\left(c_1 - n + \overline{b}_1 - h\left(y_1/w_1\right)\right) + \left(1-\overline{\pi}\right) u\left(c_2 - n + \overline{b}_2 - h\left(y_2/w_2\right)\right)\right\} \\ &\qquad + \left(1+\phi\right)\left\{\underline{\pi} u\left(c_1 - h\left(y_1/w_1\right)\right) + \left(1-\underline{\pi}\right) u\left(c_2 - h\left(y_2/w_2\right)\right)\right\} \end{aligned}$$

subject to

$$(\gamma) \qquad c_2 - h\left(y_2/w_2\right) \geq c_1 - h\left(y_1/w_2\right)$$

$$(\lambda) \qquad \begin{aligned} &\phi\left\{\overline{\pi}\left(y_1 - c_1 - \overline{b}_1\right) + \left(1-\overline{\pi}\right)\left(y_2 - c_2 - \overline{b}_2\right)\right\} \\ &\qquad + \left(1+\phi\right)\left\{\underline{\pi}\left(y_1 - c_1\right) + \left(1-\underline{\pi}\right)\left(y_2/c_2\right)\right\} = 0 \end{aligned}$$

The first order conditions on c_i and y_i, given $\overline{b}_i$, are:

$$(13) \qquad \phi\overline{\pi} u'\left(\overline{x}_1\right) + \left(1-\phi\right)\underline{\pi} u'\left(\underline{x}_1\right) - \gamma - \lambda\left[\phi\overline{\pi} + \left(1-\phi\right)\underline{\pi}\right] = 0$$

$$(14) \qquad \begin{aligned} &-\left(\phi\overline{\pi} u'\left(\overline{x}_1\right) + \left(1-\phi\right)\underline{\pi} u'\left(\underline{x}_1\right)\right) h\left(y_1/w_1\right)/w_1 + \gamma h'\left(y_1/w_2\right)/w_2 \\ &\qquad + \lambda\left[\phi\overline{\pi} + \left(1-\phi\right)\underline{\pi}\right] = 0 \end{aligned}$$

$$(15) \quad \phi\left(1-\overline{\pi}\right) u'\left(\overline{x}_2\right) + \left(1-\phi\right)\left(1-\underline{\pi}\right) u'\left(\underline{x}_2\right) + \gamma - \lambda\left[\phi\left(1-\overline{\pi}\right) + \left(1-\phi\right)\left(1-\underline{\pi}\right)\right] = 0$$

$$-\left(\phi\left(1-\overline{\pi}\right)u'\left(\overline{x}_2\right)+\left(1-\phi\right)\left(1-\underline{\pi}\right)u'\left(\underline{x}_2\right)+\gamma\right)h'\left(y_2/w_2\right)/w_2 \\ +\lambda\left[\phi\left(1-\overline{\pi}\right)+\left(1-\phi\right)\left(1-\underline{\pi}\right)\right]=0 \tag{16}$$

The form of the optimal tax structure is the standard one. For example, from (15) and (16), we obtain the zero-distortion-at-the-top result: $h'\left(y_2/w_2\right)=w_2$. Thus, y_2 is the same regardless of the size of n or $\overline{b}_i$. Let the value function for this second-stage problem be defined by $W\left(\overline{b}_1,\overline{b}_2\right)$.

In the first stage, needs credits $\overline{b}_i$ are chosen, given that they will be financed optimally. To establish whether the need credits will over- or under-compensate for needs, we investigate the welfare effects of changing the credits beginning at the situation where needs are exactly fully compensated by credits: $\overline{b}_1=\overline{b}_2=n$. Two alternative cases can be considered. In the first, the needs credit is constrained to be the same for both types, say, because it is administered separately from the income tax system. In the second case, it is allowed to differ.

Case i: $\overline{b}_1=\overline{b}_2=\overline{b}$

In the initial situation where $\overline{b}=n$ and a common tax system is applied to all households, we have $\overline{x}_1=\underline{x}_1\equiv x_1$ and $\overline{x}_2=\underline{x}_2\equiv x_2$. The envelope theorem applied to problem $\left(\mathcal{N}\right)$ yields:

$$\frac{\partial W}{\partial \overline{b}}\frac{1}{\phi}=\overline{\pi}u'\left(x_1\right)+\left(1-\overline{\pi}\right)u'\left(x_2\right)-\lambda \tag{17}$$

From first-order conditions (13) and (14), we obtain at $\overline{b}=n$:

$$\pi u'\left(x_1\right)+\left(1-\pi\right)u'\left(x_2\right)=\lambda \tag{18}$$

where π is the proportion of low-ability households in the total population. Combining (17) and (18):

$$\frac{\partial W}{\partial \overline{b}}\frac{1}{\phi}=\left(\overline{\pi}-\pi\right)\left[u'\left(x_1\right)-u'\left(x_2\right)\right]\gtrless 0 \quad \text{as} \quad \overline{\pi}\gtrless\pi$$

since $u'\left(x_1\right)>u'\left(x_2\right)$. Therefore, if there are a higher proportion of low-ability persons among the needy than among the non-needy $\left(\overline{\pi}>\underline{\pi}\right)$, needs should be over-compensated, and vice versa. Effectively, the needs credit is useful as an instrument for redistribution if needs are negatively correlated with ability[11].

11. The argument is analogous to social insurance being a useful redistributive device is the probability of ill health is correlated with ability. See the analysis in Boadway *et al.* (2006).

Case ii: $\bar{b}_1 \neq \bar{b}_2$

Suppose we again start from the case where $\bar{b}_1 = \bar{b}_2 = n$, so initially $\bar{x}_1 = \underline{x}_1 \equiv x_1$ and $\bar{x}_2 = \underline{x}_2 \equiv x_2$. Then applying the envelope theorem to problem $(\mathcal{N})$, we obtain directly:

$$\frac{\partial W}{\partial \bar{b}_1}\frac{1}{\phi} = \bar{\pi}\left[u'(x_1) - \lambda\right], \quad \frac{\partial W}{\partial \bar{b}_2}\frac{1}{\phi} = (1-\bar{\pi})\left[u'(x_2) - \lambda\right]$$

From the first-order conditions (13) and (14), we obtain at $\bar{b}_1 = \bar{b}_2 = \bar{b}$:

$$u'(x_1) = \lambda + \frac{\lambda}{\pi} > \lambda, \quad u'(x_2) = \lambda - \frac{\gamma}{1-\pi} < \lambda$$

Therefore, in this case, regardless of the distribution of ability-types by need, $\bar{b}_1$ should be increased above *n*, while $\bar{b}_2$ should be decreased below *n*. In other words, needs should be over-compensated for the low ability and under-compensated for the high ability[12].

The allocation obtained in Case ii is not identical to that obtained when two separate tax schedules can be chosen in addition to needs credits. The planner does not have enough degrees of freedom in this case to replicate the tagging solution. It might be argued that there are no good economic arguments for not applying the full tagging solution if the government is allowed to give needs credits that deviate from full compensation. This in itself could be argued to violate horizontal equity. Restricting the tax schedule to be the same for all persons may be more for administrative than economics purposes.

4.3 **Extensions**

Similar sorts of extensions could be made here as mentioned at the end of the previous section. These include, for example, multiple ability-types and multiple needs-types. As well, some other extensions suggest themselves when needs are a relevant characteristic.

First, needs may not be observable. Since needs are not observationally distinct in household behavior, households cannot be separated by need and no direct compensation can be given. At best, the progressivity of the income tax will be affected to the extent that needs are correlated with abilities (positively or negatively).

Second, tagging for needs may be imperfect. That is, there may be Type I and Type II statistical errors, as in Parsons (1996). In this case, compensation for needs will be tempered. For example, in the case in which different tax structures apply to needy and non-needy groups, the extent of compensation for needs will be reduced.

12. Rowe and Woolley (1999) had assumed that needs would be fully compensated for both types of households.

Finally, more complicated formulations of needs could be used. For example, there could be needs in leisure as well as consumption. Even if these take an additive form[13], the analysis is more complicated than our needs in consumption specification. The main reason is that compensation cannot be paid in leisure but must be paid in consumption goods. However, even if the needs in leisure is the same for both households, the value of leisure in terms of consumption goods will differ among households of different ability since the opportunity cost of taking up leisure for needs purposes is the market wage. Thus, compensation for leisure must differ across households of different ability. More discussion of this case may be found in Boadway and Pestieau (2003b).

5 Conclusions

Our ultimate purpose in this paper has been to investigate how special needs for resources — such as expenditures for medical or disability conditions — ought to be treated in the income tax system. Rowe and Woolley (1999) had suggested giving universal credit for such expenditures as part of an optimal non-linear income tax system. Our analysis suggests the matter is somewhat more complicated than that. If needs can be oberved so that households can be divided into needy and non-needy groups, full compensation for needs will be optimal if a separate tax schedule applies to the two groups. In that case, compensation for needs will be a component of the optimal inter-group lump-sum redistribution scheme. Then, the optimal tax schedule within each group will depend upon the distribution of ability types in each group. In the simple case we consider in which analytical results are possible, we find that the tax should be more progressive in the group with the highest proportion of high ability types.

On the other hand, if the government is restricted to a single economy-wide optimal tax schedule, the treatment of needs depends on the correlation of needs with ability. If a uniform credit for needs is used, it will exceed the level of needs if needs are negatively correlated with ability. This is reminiscent of results elsewhere in the social insurance literature where medical care expenditures are subsidized if the risk of ill health is inversely correlated with ability. Along the same lines, Cremer *et al* (2001) study optimal income design in a setting in which households differ in both ability and an unobserved endowment. The endowment plays a very similar role to needs here: persons with a smaller endowment of a commodity have larger ‘needs’ for the commodity. They show how differential commodity tax system can improve welfare if used alongside the income tax in a setting in which the well-known Atkinson-Stiglitz (1976) theorem otherwise applies. They also find that the self-selection constraints can take very surprising patterns. If, however, needs can be differentiated by income level, needs will be over-compensated for low-ability households, and under-compensated for high-ability households, regardless of the correlation between needs and ability.

13. Thus, preferences take the form $c-h(\ell+n)$, where n are needs in leisure.

The results have some potential policy implications. The simplest one is that whenever the population can be divided into identifiable groups with different ability distributions, it is optimal to have both separate income tax schedules in the different groups and lump-sum redistributive transfers from groups with a higher proportion of high-ability types to those with a lower proportion. Moreover, the inter-group transfers should be augmented to the extent that needs differences vary between groups. One obvious example of this is a federation with regions of different ability distributions. It is seemingly efficient to decentralize redistribution to the regional level, and restrict the federal government to inter-regional equalization transfers based on ability levels, and if necessary, needs levels. Of course, there will be two caveats to this prescription. First, if horizontal equity must be respected, it might be necessary to impose a uniform tax schedule nationwide. And, second, if households are mobile, so that ability distributions across regions are endogenous, things become much more complicated[14].

More generally, the analysis informs the issue of how costs for health and disability expenditures ought to be accounted for in the income tax system. Assuming needs for these expenditures can be observed, they can be used as part of a system of redistributing not only in favor of high-need persons, but also in favor of low-ability persons.

References

AKERLOF G.A. (1978). – "The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Training", *American Economic Review* 68, pp. 8–19.

ATKINSON A.B. and STIGLITZ J.E. (1976). – "The Design of Tax Structure: Direct versus Indirect Taxation", *Journal of Public Economics* 6, pp. 55-75.

BLACKORBY C. and DONALDSON D. (1994). – "Information and Intergroup Transfers", *American Economic Review Papers and Proceedings* 84, pp. 440-47.

BOADWAY R. (2004). – "The Theory and Practice of Equalization", *CESifo Economic Studies* 50, pp. 211-54.

BOADWAY R., LEITE-MONTEIRO M., MARCHAND M. and PESTIEAU P. (2006). – "Social Insurance and Redistribution with Moral Hazard and Adverse Selection", *Scandinavian Journal of Economics* 108, pp. 279-98.

BOADWAY R. and PESTIEAU P. (2003a). – "Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem", in Richard Arnott, Bruce Greenwald, Ravi Kanbur and Barry Nalebuff (eds.), *Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz* (Cambridge, Mass: MIT Press), pp. 387-403.

BOADWAY R. and PESTIEAU P. (2003b). – "Optimal Redistribution and Needs", in Michael Ahlheim, Heinz-Dieter Wenzel and Wolfgang Weigard (eds.), *Steuerpolitik-Von der Theorie zur Praxis*, Festschrift fhr Manfred Rose (Berlin: Springer-Verlag), pp. 1-14.

BOSSERT W. and WEYMARK J.A. (2004). – "Utility in Social Choice", in S. Barberà, P.J. Hammond and C. Seidl (eds.), *Handbook of Utility Theory, Volume 2: Extensions* (Boston, Dordrecht and London: Kluwer Academic Publishers), pp. 1099-177.

BRETT C. and WEYMARK J.A. (2004). – "Public Good Provision and the Comparative Statics of Optimal Nonlinear Income Taxation", *International Economic Review,* forthcoming.

CREMER H., PESTIEAU P. and ROCHET J.-C. (2001). – "Direct Versus Indirect Taxation: The Design of the Tax Structure Revisited", *International Economic Review* 42, pp. 781-99.

DIAMOND P.A. (1998). – "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates", *American Economic Review* 88, pp. 83-95.

14. For a general discussion of these issues, see Boadway (2004).

GUESNERIE R. and SEADE J. (1982). – "Nonlinear Pricing in a Finite Economy", *Journal of Public Economics* 17, pp. 157-180.

HAMILTON P. and PESTIEAU P. (2005). – "Optimal Income Taxation and the Ability Distribution: Implications for Migration Equilibria", *International Tax and Public Finance* 12, pp. 29-45.

HERNANZ V., MALHERBET F. and PELLIZZARI M. (2004). – "Take-Up of Welfare Benefits in OECD Countries: A Review of the Evidence", OECD Social, Employment and Migration Working Paper No. 17.

IMMONEN R., KANBUR R., KEEN M. and TUOMALA M. (1998). – "Tagging and Taxing: The Optimal Use of Categorical and Income Information in Designing Tax/Transfer Schemes", *Economica* 65, pp. 179-92.

JACQUET L. and VAN DER LINDEN B. (2006). – "The Normative Analysis of Tagging Revisited: Dealing with Stigmatization", *FinanzArchiv: Public Finance Analysis* 62, pp. 168-98.

KEEN M. (1992). – "Needs and Targeting", *Economic Journal* 102, pp. 67-79.

KOLM S.C. (1969). – "The Optimal Production of Social Justice", in J. Margolis and H. Guitton (eds.), *Public Economics* (London: Macmillan), pp. 145-200.

MIRRLEES J.A. (1971). – "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies* 38, pp. 175-208.

MIRRLEES J.A. (1974). – "Notes on Welfare Economics, Information, and Uncertainty", in M. Balch, D. McFadden and S. Wu (eds.), *Essays on Equilibrium Behavior under Uncertainty* (Amsterdam, North-Holland) pp. 243-58.

PARSONS D.O. (1996). – "Imperfect 'Tagging' in Social Insurance Programs", *Journal of Public Economics* 62, pp. 183-207.

POLLAK R.A. (1971). – "Additive Utility Functions and Linear Engel Curves", *Review of Economic Studies* 38, pp. 401-14.

ROWE N. and WOOLLEY F. (1999). – "The Efficiency Case for Universality", *Canadian Journal of Economics* 32, pp. 613-29.

SALANIÉ B. (2002). – "Optimal Demogrants with Imperfect Tagging", *Economic Letters* 75, pp. 319-24.

STIGLITZ J.E., (1982). – "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics* 17, pp. 213-40.

VIARD A.D. (2001a). – "Optimal Categorical Transfer Payments: The Welfare Economics of Limited Lump-Sum Redistribution", *Journal of Public Economic Theory* 3, pp. 483-500.

VIARD A.D. (2001b). – "Some Results on the Comparative Statics of Optimal Categorical Transfer Payments", *Public Finance Review* 29, pp. 148-80.

WEYMARK J.A. (1987). – "Comparative Static Properties of Optimal Nonlinear Income Taxes", *Econometrica* 55, pp. 1165-85.